PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/15, 15/58, 15/62
C12N 9/70, C07K 7/10
C12N 5/10, 1/19, 1/21
A61K 37/64, 37/54

A1

(11) International Publication Number: WO 91/09125

(43) International Publication Date: 27 June 1991 (27.06.91)

(21) International Application Number: PCT/GB90/01911

(22) International Filing Date: 7 December 1990 (07.12.90)

(30) Priority data:
8927722.2 7 December 1989 (07.12.89) GB

(71) Applicant *(for all designated States except US)*: BRITISH BIO-TECHNOLOGY LIMITED [GB/GB]; Watlington Road, Cowley, Oxford OX4 5LY (GB).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)* : DAWSON, Keith, Martyn [GB/GB]; 80 Barnards Hill, Marlow, Bucks SL7 2NZ (GB). HUNTER, Michael, George [GB/GB]; 7 Nash Close, Aylesbury, Bucks HP21 7YB (GB). CZAPLEWSKI, Lloyd, George [GB/GB]; No. 3 Merton Close, Didcot OX11 8UJ (GB).

(74) Agents: SHEARD, Andrew, Gregory et al.; Kilburn & Strode, 30 John Street, London WC1N 2DD (GB).

(81) Designated States: AT (European patent), AU, BE (European patent), CA, CH (European patent), DE (European patent), DK (European patent), ES (European patent), FI, FR (European patent), GB (European patent), GR (European patent), HU, IT (European patent), JP, KR, LU (European patent), NL (European patent), NO, SE (European patent), US.

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: PROTEINS AND NUCLEIC ACIDS

(57) Abstract

Relatively inactive fusion proteins are activatable by enzymes of the clotting cascade to have fibrinolytic and/or clot formation inhibition activity. For example, a fusion protein comprising two hirudin or streptokinase molecules, linked by a cleavable linkage sequence, may be cleaved to yield anti-thrombotic hirudin or fibrinolytic streptokinase by thrombin or Factor Xa. Fibrinolytic or clot formation inhibition activity is therefore directed to the site of clot formation.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria	FI	Finland	ML	Mali
AU	Australia	FR	France	MN	Mongolia
BB	Barbados	GA	Gabon	MR	Mauritania
BE	Belgium	GB	United Kingdom	MW	Malawi
BF	Burkina Faso	GN	Guinea	NL	Netherlands
BG	Bulgaria	GR	Greece	NO	Norway
BJ	Benin	HU	Hungary	PL	Poland
BR	Brazil	IT	Italy	RO	Romania
CA	Canada	JP	Japan	SD	Sudan
CF	Central African Republic	KP	Democratic People's Republic of Korea	SE	Sweden
CG	Congo			SN	Senegal
CH	Switzerland	KR	Republic of Korea	SU	Soviet Union
CI	Côte d'Ivoire	LI	Liechtenstein	TD	Chad
CM	Cameroon	LK	Sri Lanka	TG	Togo
DE	Germany	LU	Luxembourg	US	United States of America
DK	Denmark	MC	Monaco		
ES	Spain	MG	Madagascar		

PROTEINS AND NUCLEIC ACIDS

This invention relates to proteinaceous compounds which can be cleaved to release fibrinolytic and/or anti-thrombotic activity. It also relates to nucleic acid (DNA and RNA) coding for all or part of such compounds. In preferred embodiments, the invention relates to fusion proteins produced by linking together fibrinolytic and/or anti-thrombotic proteins with a cleavable linker, their preparation, pharmaceutical compositions containing them and their use in the treatment of thrombotic disease.

The fibrinolytic system is the natural counterpart to the clotting system in the blood. In the process of blood coagulation, a cascade of enzyme activities are involved in generating a fibrin network which forms the framework of a clot, or thrombus. Degradation of the fibrin network (fibrinolysis) is accomplished by the action of the enzyme plasmin. Plasminogen is the inactive precursor of plasmin and conversion of plasminogen to plasmin is accomplished by cleavage of the peptide bond between arginine 561 and valine 562 of plasminogen. Under physiological conditions this cleavage is catalysed by tissue-type plasminogen activator (tPA) or by urokinase-type plasminogen activator (uPA).

If the balance between the clotting and fibrinolytic systems becomes locally disturbed, intravascular clots may form at inappropriate locations leading to conditions such as coronary thrombosis and myocardial infarction, deep vein thrombosis, stroke, peripheral

arterial occlusion and embolism. In such cases, the administration of fibrinolytic and anti-thrombotic agents has been shown to be a beneficial therapy for the promotion of clot dissolution.

Fibrinolytic therapy has become relatively widespread with the availability of a number of plasminogen activators such as tPA, uPA, streptokinase and the anisoylated plasminogen streptokinase activator complex, APSAC. Each of these agents has been shown to promote clot lysis, but all have deficiencies in their activity profile which makes them less than ideal as therapeutic agents for the treatment of thrombosis (reviewed by Marder and Sherry, New England Journal of Medicine 1989, **318**: 1513-1520).

A major problem shared by all of these agents is that at clinically useful doses, they are not thrombus specific as they activate plasminogen in the general circulation. The principal consequence of this is that proteins such as fibrinogen involved in blood clotting are destroyed and dangerous bleeding can occur. This also occurs with tPA despite the fact that, at physiological concentrations, it binds to fibrin and shows fibrin selective plasminogen activation.

Another important shortcoming in the performance of existing plasminogen activators is that re-occlusion of the reperfused blood vessel commonly occurs after cessation of administration of the thrombolytic agent. This is thought to be due to the persistence of thrombogenic material at the site of thrombus dissolution.

Anti-thrombotic proteins may be used in the treatment or prophylaxis of thrombosis either alone or as an adjunct to fibrinolytic agents. Suitable anti-thrombotic proteins include hirudin, activated protein C and anti-thrombin III.

An alternative approach to enhancing fibrinolysis and inhibition of blood clotting has now been devised which is based on the use of fusion proteins cleavable to achieve release of fibrinolytic and/or anti-thrombotic activity at the site of blood clotting. To achieve this, proteins involved in fibrinolysis or inhibition of coagulation are joined by a linker region which is cleavable by an enzyme involved in blood clotting. Examples of proteins which may be incorporated into such a cleavable protein include tPA, uPA, streptokinase, plasminogen, activated protein C, hirudin and anti-thrombin III. Fusion of such proteins to a protein with a favourable property not directly related to dissolution of blood clots, for example albumin which has a long plasma half-life, may also be beneficial. An advantage of this approach is that thrombus selectivity of fibrinolytic or inhibition of clot formation activity is achieved by way of the thrombus-specific localisation of the cleaving enzymes.

According to a first aspect of the invention, there is provided a fusion protein comprising a first sequence and a second sequence, the fusion protein being cleavable between the first and second sequences by an enzyme involved in blood clotting, wherein after the fusion protein is so cleaved the first and second sequences, or either of them, has greater fibrinolytic

and/or anti-thrombotic activity than the uncleaved fusion protein.

The fusion protein may be a cleavable dimer of two fibrinolytic and/or anti-thrombotic proteins, such as hirudin or streptokinase. It may be a homodimer or a heterodimer. The fusion protein may have substantially reduced or no fibrinolytic and/or anti-thrombotic activity compared to the cleavage products, but a certain amount of activity in the fusion protein can be tolerated. It is not necessary for both the cleavage products to have fibrinolytic and/or anti-thrombotic activity, but it is preferred for them to do so.

The fusion protein is not restricted to being a dimer; it may have any number (such as three, four or more) sequences which are cleavable one from the other, compatible with the therapeutic utility of the protein. At least one, and preferably more than one or even all, of the sequences resulting from the cleavage will have greater activity than the fusion protein, or a combination of some or all of the cleavage products will collectively have such greater activity. In any event, cleavage will result in a net increase in or release of activity.

Proteinaceous compounds in accordance with the first aspect of the invention, are therefore cleaved to release activity in at least one of two ways. First, a compound may be cleaved to release fibrinolytic activity. Secondly, a compound may be cleaved to release anti-thrombotic activity. Conceivably, a compound may be cleaved to release both functions. It

should be noted that a released fragment of the fusion protein may have fibrinolytic activity directly (in that it lyses fibrin) or indirectly (in that it causes activation of a molecule which leads to lysis of fibrin).

One preferred proteinaceous compound which is cleavable to have enhanced anti-thrombotic activity is a fusion protein of two hirudin molecules linked (for example carboxy terminus to amino terminus) by a linker amino acid sequence cleavable, for example, by Factor Xa.

Hirudins are naturally occurring polypeptides of 65 or 66 amino acids in length that are produced by the leech Hirudo medicinalis. Hirudin is an anticoagulating agent which binds to thrombin and prevents blood coagulation by inhibiting thrombin from catalysing the conversion of fibrinogen to fibrin, thus preventing the formation of the protein framework of blood clots. The binding of hirudin also prevents other prothrombic activities of thrombin including activation of factors V, VII, XIII and platelets. There are three principal variants of hirudin (named HV-1, HV-2 and HV-3).

Another preferred fusion protein comprises two streptokinase molecules linked (for example carboxy terminus to amino terminus) by a linker amino acid sequence cleavable, for example, by thrombin.

Streptokinase is a 414 amino acid, 47kDa protein secreted by many pathogenic streptococci of different serogroups. It is a plasminogen activator but, unlike mammalian plasminogen activators, it is not a protease

and it activates plasminogen by forming a binary complex with plasminogen (SK-plasminogen) which functions as an activator of free plasminogen. Streptokinase is effective in inducing clot lysis in the treatment of myocardial infarction and is widely used for this indication.

Cleavable fusion proteins within the scope of this invention may have reduced fibrinolytic and/or anti-thrombotic activity compared to their component molecules; cleavage releases the component molecules which possess to an adequate degree the activity of their wild-type parent molecules.

The blood coagulation mechanism comprises a series of enzyme reactions which culminate in the production of insoluble fibrin, which forms the mesh-like protein framework of blood clots. Thrombin is the enzyme responsible for the conversion of soluble fibrinogen to fibrin. Conversion of prothrombin, the inactive precursor of thrombin, to thrombin is catalysed by activated Factor X (Factor Xa). (Thrombin is also known as Factor IIa, and prothrombin as Factor II.)

Factor Xa is generated from Factor X extrinsically or intrinsically. In the extrinsic route, Factor VII is activated to Factor VIIa, which generates Factor Xa from Factor X. In the intrinsic route, the activation of Factor X to Factor Xa is catalysed by Factor IXa. Factor IXa is generated from Factor IX by the action of Factor XIa, which in turn is generated by the action of Factor XIIa on Factor XI. Factor XIIa is generated from Factor XII by the action of Kallikrein. Factors

VIIIa and Va are thought to act as cofactors in the activation of Factors X and II, respectively.

Fibrin, as first formed from fibrinogen, is in the loose form. Loose fibrin is converted to tight fibrin by the action of Factor XIIIa, which crosslinks fibrin molecules.

Activated protein C is an anticoagulant serine protease generated in the area of clot formation by the action of thrombin, in combination with thrombomodulin, on protein C. Activated protein C regulates clot formation by cleaving and inactivating the pro-coagulant cofactors Va and VIIIa.

The term "enzyme involved in blood clotting" as used in this specification therefore includes kallikrein Factors XIIa, XIa, IXa, VIIa, Xa and thrombin (Factor IIa), which are directly involved in the formation of fibrin and activated protein C, which is involved in the control of blood clotting. The most preferred enzymes are Factor Xa and thrombin because they are most immediately involved with fibrin formation.

Generation and activity of at least Factor Xa and thrombin is tightly regulated to ensure that thrombus generation is restricted to the site of the thrombogenic stimulus. This localisation is achieved by the combined operation of at least two control mechanisms: the blood clotting enzymes function as complexes intimately associated with the phospholipid cellular membranes of platelets and endothelial cells at the site of vascular injury (Mann, K. G., 1984, in:

"Progress in Hemostasis and Thrombosis", 1 - 24, ed Spaet, T. H. Grune and Stratton); and, free thrombin or Factor Xa released from the thrombus site into the circulation is rapidly inactivated by the action of proteinase inhibitors such as anti-thrombin III.

Thus, the activity of the penultimate (Factor Xa) and the final (thrombin) enzymes in the clotting cascade are particularly well localised to the site of thrombus generation and for this reason are preferred.
Thrombin has been found to remain associated with thrombi and to bind non-covalently to fibrin. On digestion of thrombi with plasmin, active thrombin is liberated and is thought to contribute to the reformation of thrombi and the re-occlusion of vessels which commonly occurs following thrombolytic treatment with plasminogen activators (Bloom A. L., 1962, Br. J. Haematol, 82, 129; Francis et al, 1983, J. Lab. Clin. Med., 102, 220; Mirshahi et al, 1989, Blood 74, 1025).

For these reasons, it is preferred in certain embodiments of the invention to produce fusion proteins activatable by thrombin or Factor Xa thereby to create a preferred class of thrombus-selective, fibrinolytic proteins. The most preferred of these fusion proteins regain the favourable properties of the parent molecules upon cleavage and exhibit thrombus selectivity by the novel property of being cleaved to release the component proteins of the fusion protein at the site of new thrombus formation by the action of one of the enzymes involved in generation of the thrombus and preferably localised there.

Factor Xa (E.C.3.4.21.6) is a serine protease which converts human prothrombin to thrombin by specific cleavage of the Arg(273)-Thr(274) and Arg(322)-Ile(323) peptide bonds (Mann et al 1981, Methods in Enzymology **80** 286-302). In human prothrombin, the Arg(273)-Thr(274) site is preceded by the tripeptide Ile-Glu-Gly and the Arg(322)-Ile(323) site is preceded by the tripeptide Ile-Asp-Gly. The structure required for recognition by Factor Xa appears to be determined by the local amino acid sequence preceding the cleavage site (Magnusson et al, 1975, in: "Proteases and Biological Control", 123-149, eds., Reich et al, Cold Spring Harbor Laboratory, New York). Specificity for the Ile-Glu-Gly-Arg and Ile-Asp-Gly-Arg sequence is not absolute as Factor Xa has been found to cleave other proteins, for example Factor VIII at positions 336, 372, 1689 and 1721, where the preceding amino acid sequence differs significantly from this format (Eaton et al, 1986 Biochemistry **25** 505-512). As the principal natural substrate for Factor Xa is prothrombin, preferred recognition sequences are those in which arginine and glycine occupy the P1 and P2 positions, respectively, an acidic residue (aspartic or glutamic acid) occupies the P3 position and isoleucine or another small hydrophobic residue (such as alanine, valine, leucine or methionine) occupies the P4 position. However, as Factor Xa can cleave sequences which differ from this format, other sequences cleavable by Factor Xa may be used in the invention, as can other sequences cleavable by other enzymes of the clotting cascade.

In order to make fusion proteins which are cleavable by these preferred enzymes, the amino acid sequence linking the components of the fusion protein must be recognised as a cleavage site for these preferred enzymes. To make fusion proteins which are cleaved by, for example, Factor Xa, an amino acid sequence cleavable by Factor Xa may be used to link the two components (that is, the first and second, and possibly other, sequences) of the fusion protein. The sequence Ile-Glu-Gly-Arg which is at one of the sites in prothrombin cleaved by Factor Xa may be such a sequence. Other possibilities would be sequences or mimics of sequences cleaved by Factor Xa in other proteins or peptides. DNA coding for the Ile-Glu-Gly-Arg sequence as the carboxy-terminal part of a cleavable linker as a protein production aid is disclosed in UK Patent Application GB-A-2160206 but the use of an Ile-Glu-Gly-Arg sequence for the purpose of this invention is not disclosed in that specification.

Cleavage of fusion proteins by an enzyme of the clotting cascade such as thrombin or Factor Xa can be measured in a number of ways, for example by SDS-PAGE analysis, and by assaying for the functions of one or more of the cleavage products of the fusion protein.

Thrombin (E.C. 3.4.21.5) is a serine protease which catalyses the proteolysis of a number of proteins including fibrinogen (A alpha and B beta chains), Factor XIII, Factor V, Factor VII, Factor VIII, protein C and anti-thrombin III. The structure required for recognition by thrombin appears to be partially determined by the local amino acid sequence around the

cleavage site but is also determined to a variable extent by sequence(s) remote from the cleavage site. For example, in the fibrinogen A alpha chain, residues P2 (Val), P9 (Phe) and P10 (Asp) are crucial for α-thrombin-catalysed cleavage at the Arg(16)-Gly(17) peptide bond (Ni, F. et al 1989, Biochemistry 28 3082-3094). Comparative studies of several proteins and peptides which are cleaved by thrombin has led to the proposal that optimum cleavage sites for α-thrombin may have the structure of (i) P4-P3-Pro-Arg-P1'-P2', where each of P3 and P4 is independently a hydrophobic amino acid (such as valine) and each of P1' and P2' is independently a non-acidic amino acids, or (ii) P2-Arg-P1' where P2 or P1' is glycine (Chang, J. 1985, Eur. J. Biochem. 151 217-224). There are, however, exceptions to these general structures which are cleaved by thrombin and which may be used in the invention.

To produce a fusion protein which could be cleaved by thrombin, a linker sequence containing a site recognised and cleaved by thrombin may be used. An amino acid sequence such as that cleaved by thrombin in the fibrinogen A alpha chain may be used. Other possible sequences would include those involved in the cleavage by thrombin of fibrinogen B beta, Factor XIII, Factor V, Factor VII, Factor VIII, protein C, anti-thrombin III and other proteins whose cleavage is catalysed by thrombin. An example of a thrombin cleavable linker may be the sequence Gly-Pro-Arg which is identical to that found at positions 17-20 in fibrinogen A alpha. This is not the principal thrombin cleavage site in fibrinogen A alpha but thrombin can

cleave the Arg(19)-Val(20) peptide bond. Another suitable thrombin cleavable linker sequence is Val-Glu-Leu-Gln-Gly-Val-Val-Pro-Arg which is identical to that found in Factor XIII.

In a preferred embodiment the invention relates to fusion proteins of streptokinase and/or hirudin linked by peptide sequences which are cleaved by thrombin, Factor Xa or other enzymes involved in blood clotting to release products with fibrinolytic and/or anti-thrombotic activity.

Fusion proteins in accordance with the invention may contain other modifications (as compared to wild-type counterparts of their components such as streptokinase and hirudin) which may be one or more additions, deletions or substitutions. An example of such a modification would be streptokinase variants in which inappropriate glycosylation during yeast expression was prevented by substitution of sequences recognised as glycosylation signals by yeast. Another example would be the addition of an Arg-Gly-Asp-Xaa sequence, where Xaa represents a variable amino acid such as Ser, to the carboxy terminus of the fusion to enhance its plasma lifetime.

Preferred features of fusion proteins within the scope of the invention also apply, where appropriate, to other compounds of the invention, mutatis mutandis.

Fusion proteins in accordance with the first aspect of the invention can be synthesised by any convenient route. According to a second aspect of the invention

there is provided a process for the preparation of a proteinaceous compound as described above, the process comprising coupling successive amino acid residues together and/or ligating oligopeptides. Although proteins may in principle be synthesised wholly or partly by chemical means, the route of choice will be ribosomal translation, preferably in vivo, of a corresponding nucleic acid sequence. The protein may be glycosylated appropriately.

It is preferred to produce proteins in accordance with the invention by using recombinant DNA technology. DNA encoding each of the first and second sequences of the fusion protein may be from a cDNA or genomic clone or may be synthesised. Amino acid substitutions, additions or deletions are preferably introduced by site-specific mutagenesis. Suitable DNA sequences encoding streptokinase and hirudin and other polypeptide sequences useful in the scope of the invention may be obtained by procedures familiar to those having ordinary skill in genetic engineering. For several proteins, it is a routine procedure to obtain recombinant protein by inserting the coding sequence into an expression vector and transfecting or transforming the vector into a suitable host cell. A suitable host may be a bacterium such as E. coli, a eukaryotic microorganism such as yeast or a higher eukaryotic cell.

According to a third aspect of the invention, there is provided synthetic or recombinant nucleic acid coding for a proteinaceous compound as described above. The nucleic acid may be RNA or DNA. Preferred

characteristics of this aspect of the invention are as for the first aspect.

According to a fourth aspect of the invention, there is provided a process for the preparation of nucleic acid in accordance with the third aspect, the process comprising coupling successive nucleotides together and/or ligating oligo- and/or polynucleotides.

Recombinant nucleic acid in accordance with the third aspect of the invention may be in the form of a vector, which may for example be a plasmid, cosmid or phage. The vector may be adapted to transfect or transform prokaryotic (for example bacterial) cells and/or eukaryotic (for example yeast or mammalian) cells. A vector will comprise a cloning site and usually at least one marker gene. An expression vector will have a promoter operatively linked to the sequence to be inserted into the cloning site and, preferably, a sequence enabling the protein product to be secreted. Expression vectors and cloning vectors (which need not be capable of expression) are included in the scope of the invention.

It is to be understood that the term "vector" is used in this specification in a functional sense and is not to be construed as necessarily being limited to a single nucleic acid molecule.

Using a vector, for example as described above, fusion proteins in accordance with the invention may be expressed and secreted into the cell culture medium in a biologically active form without the need for any

additional biological or chemical procedures. Suitable cells or cell lines to be transformed may be mammalian cells which grow in continuous culture and which can be transfected or otherwise transformed by standard techniques. Examples of suitable cells include Chinese hamster ovary (CHO) cells, mouse myeloma cell lines such as P3X63-Ag8.653, COS cells, HeLa cells, BHK cells, melanoma cell lines such as the Bowes cell line, mouse L cells, human hepatoma cell lines such as Hep G2, mouse fibroblasts and mouse NIH 3T3 cells. Such cells may be particularly appropriate for expression when one or more of the protein sequences constituting the fusion protein is of mammalian derivation, such as tissue plasminogen activator (t-PA).

Yeast (for example Pichia pastoris or Saccharomyces cerevisiae) or bacteria (for example Escherichia coli) may be preferred for the expression of many of the fusion proteins of the invention, as may insect cells such as those which are Baculovirus-infected.

Compounds of the present invention may be used within pharmaceutical compositions for the prevention or treatment of thrombosis or other conditions where it is desired to produce local fibrinolytic and/or anticoagulant activity. Such conditions include myocardial and cerebral infarction, arterial and venous thrombosis, thromboembolism, post-surgical adhesions, thrombophlebitis and diabetic vasculopathies.

According to a fifth aspect of the invention, there is provided a pharmaceutical composition comprising one or more compounds in accordance with the first aspect of

the invention and a pharmaceutically or veterinarily acceptable carrier. Such a composition may be adapted for intravenous administration and may thus be sterile. Examples of compositions in accordance with the invention include preparations of sterile fusion proteins in isotonic physiological saline and/or buffer. The composition may include a local anaesthetic to alleviate the pain of injection. Compounds of the invention may be supplied in unit dosage form, for example as a dry powder or water-free concentrate in a hermetically sealed container such as an ampoule or sachet indicating the quantity of protein. Where a compound is to be administered by infusion, it may be dispensed by means of an infusion bottle containing sterile water for injections or saline or a suitable buffer. Where it is to be administered by injections, it may be dispensed with an ampoule of water for injection, saline or a suitable buffer. The infusible or injectable composition may be made up by mixing the ingredients prior to administration. Where it is to be administered as a topical treatment, it may be dispensed in a suitable base.

The quantity of material to be administered will depend on the amount of fibrinolysis or inhibition of clotting required, the required speed of action, the seriousness of the thromboembolic position and the size of the clot. The precise dose to be administered will, because of the very nature of the condition which compounds of the invention are intended to treat, be determined by the physician. As a guideline, however, a patient being treated for a mature thrombus will generally

receive a daily dose of a fusion protein of from 0.01 to 10 mg/kg of body weight either by injection in for example up to 5 doses or by infusion.

The invention may be used in a method for the treatment or prophylaxis of thombosis, comprising the administration of an effective non-toxic amount of a compound in accordance with the first aspect. According to a further aspect of the invention, there is therefore provided the use of a compound as described above in the preparation of a thombolytic and/or anticoagulant agent.

The invention concerns especially the DNAs, the vectors, the transformed host strains, the fusion proteins and the process for the preparation thereof as described in the examples.

The following examples of the invention are offered by way of illustration, and not by way of limitation. The examples refer to the accompanying drawings, in which:

Figure 1 shows schematically the arrangement of a set of oligonucleotides used in the assembly of a synthetic hirudin gene (Preparation 1);

Figure 2 shows a map of plasmid pSW6 (Preparation 2);

Figure 3 shows a map of plasmid pJK1 (Preparation 2);

Figure 4 shows a map of plasmid pGC517 (Example 4);

Figure 5 shows a zymograph of E. coli strains expressing streptokinase activity (Example 11); and

Figure 6 shows a zymograph demonstrating cleavage of a streptokinase-streptokinase fusion protein by thrombin (Example 13).

Methodology

The techniques of genetic engineering and genetic manipulation used in the manufacture of the genes described and in their further manipulation for construction of expression vectors are well known to those skilled in the art. Descriptions of modern techniques can be found in the laboratory manuals "Current Protocols in Molecular Biology" , Volumes 7 and 2, edited by F. M. Ausubel et al, published by Wiley-Interscience, New York and in "Molecular Cloning, A Laboratory Manual" (second edition) edited by Sambrook, Fritsch and Maniatis published by Cold Spring Harbor Laboratories, New York. M13mp18, M13mp19 and pUC19 DNAs were purchased from Pharmacia Ltd., Midsummer Boulevard, Central Milton Keynes, Bucks, MK9 3HP, United Kingdom. Restriction endonucleases were purchased either from Northumbria Biologicals Limited, South Nelson Industrial Estate, Cramlington, Northumberland, NE23 9HL, United Kingdom or from New England Biolabs, 32 Tozer Road, Beverly, MA 01915-5510 USA. E. coli HW1110 (lacIq) is used as expression host

in certain of the following examples: a suitable commercially available alternative is JM109, available from Northumbria Biologicals Ltd.

PREPARATION 1 - Construction of a Hirudin HV1 gene

A. Gene Design

A synthetic hirudin HV-1 gene was designed based on the published amino acid sequence (Dodt J., et al FEBS Letters **165** 180 (1984)). Unique restriction endonuclease target sites were incorporated to facilitate subsequent genetic manipulation (see SEQ. ID NO:1 in the Sequence Listings immediately before the claims). The codons selected were those favoured by either S. cerevisiae or E. coli and are thus suitable for expression in either organism.

B. Gene Construction

The gene sequence was divided into 12 oligodeoxyribonucleotides (see SEQ. ID NO:2) such that after annealing each complementary pair 2 oligonucleotides, they were left with cohesive ends either for or of 7 bases in length.

C. Oligonucleotide Synthesis

The oligonucleotides were synthesised by automated phosphoramidite chemistry on an Applied Bio-Systems 380B DNA Synthesiser, using cyanoethyl phosphoramidites. The methodology is now widely used and has already been described (Beaucage, S.L. and

Caruthers, M.H. Tetrahedron Letters **24**, 245 (1981) and Caruthers, M. H. Science **230**, 281-285 (1985)).

D. Gene Assembly

The oligonucleotides were kinased to provide them with a 5′ phosphate to allow their subsequent ligation. The oligonucleotides were assembled as shown in Figure 1.

Kinasing of Oligomers

100 pmole of oligomer was dried down and resuspended in 20 µl kinase buffer (70 mM Tris, pH 7.6, 10 mM $MgCl_2$, 1 mM ATP, 0.2 mM spermidine, 0.5 mM dithiothreitol (DTT)). T4 polynucleotide kinase (2 mcl. 10 000 U/ml) was added and the mixture was incubated at 37°C for 30 minutes. The kinase was then inactivated by heating at 70°C for 10 minutes.

Complementary pairs of kinased oligonucleotides were annealed in pairs (90°C, 5 minutes, followed by slow cooling at room temperature). The 6 paired oligomers were then mixed together, incubated at 50°C for 5 minutes and allowed to cool. They were then ligated overnight at 16°C with T4 DNA ligase. The strategy is shown diagrammatically in Figure 1 (note P = 5′-phosphate). To prevent possible multimerisation, oligomers designated BB2011 and BB2020 were not kinased. The sequences of the oligomers shown in Figure 1 correspond to those given in SEQ.ID NO:2.

The ligation products were separated on a 2% low gelling temperature agarose gel and the DNA fragment of

ca. 223 base pairs corresponding to the hirudin HV-1 gene was excised and extracted from the gel. The purified fragment was then ligated to HindIII and EcoRI treated pUC19 plasmid DNA. The transformation of E. coli host strains was accomplished using standard procedures. The strain used as a recipient in the transformation of plasmid vectors was HW87 which has the following genotype:

araD139(ara-leu)DELTA7697 (lacIPOZY)DELTA74 galU

galK hsdR rpsL srl recA56

The use of HW87 was not critical: any suitable recipient strain could be used, for example, E. coli AG1, which is available from Northumbria Biologicals Ltd. The recombinant ligation products were transformed into E. coli K12 host strain HW87 and plated onto Luria-agar ampicillin (100 μg/ml) plates. Twelve ampicillin-resistant colonies were picked and used to prepare plasmid DNA for sequence analysis. Double stranded dideoxy sequence analysis using a universal sequencing primer BB22 (5'-CAGGGTTTTCCCAGTCACG-3'), (SEQ ID NO:3) complementary to the universal primer region of pUC19 was used to identify a correct clone pUC19 HV-1. The pUC19 recombinant was used to construct an expression vector.

PREPARATION 2 - Construction of a Hirudin HV1 Expression Vector

An expression vector was designed to enable the secretion of hirudin to the extracellular medium after expression in S. cerevisiae. Secretion of hirudin is desirable as this facilitates production of the protein with an authentic N-terminus. It also eases purification, limits intracellular proteolysis, reduces potential toxic effects on the yeast host and allows optimal protein folding and formation of native disulphide bonds. Secretion of hirudin through the yeast membrane was directed by fusion of hirudin to the yeast mating type alpha-factor pre-pro-peptide (a naturally secreted yeast peptide).

The yeast expression vector pSW6 (Figure 2) is based on the 2 μ circle from S. cerevisiae. (pSW6 was deposited in S. cerevisiae strain BJ2168 at The National Collection of Industrial and Marine Bacteria Limited, 23 St. Machar Drive, Aberdeen, AB2 1RY, Scotland, United Kingdom on 23rd October 1990 under Accession No. NCIMB 40326.) pSW6 is a shuttle vector capable of replication in both E. coli and S. cerevisiae and contains an origin of DNA replication for both organisms, the leu2 gene (a selectable marker for plasmid maintenance in the yeast host) and the ampicillin resistant locus for selection of plasmid maintenance in E. coli. (The DNA sequence of the vector has been determined; the E. coli sequences are derived from the E. coli ColE1-based replicon pAT153.) The full sequence is given as SEQ.ID:4. The ability to passage this vector through E. coli greatly

facilitates its genetic manipulation and ease of purification. pSW6 contains an α-factor pre-pro-peptide gene fused in-frame to the gene for epidermal growth factor (EGF). The expression of this fusion is under the control of an efficient galactose regulated promoter which contains hybrid DNA sequences from the S. cerevisiae GAL 1-10 promoter and the S. cerevisiae phosphoglycerate kinase (PGK) promoter. Transcription of the EGF gene is terminated in this vector by the natural yeast PGK terminator. The EGF gene in pSW6 can be removed by digestion with restriction endonucleases HindIII and BamHI. This removes DNA encoding both EGF and 5 amino acids from the C-terminus of the α-factor pro-peptide. Genes to be inserted into the pSW6 expression vector must therefore have the general composition: HindIII site - α-factor adaptor - gene- BamHI site.

To rebuild the DNA encoding the amino acids at the C-terminal end of the α-factor pro-peptide and to fuse this to the synthetic hirudin gene, an oligonucleotide adapter (5'-AGCTTGGATAAAAGA-3' (top strand, SEQ.ID:5), 5'-TCTTTTATCCA-3' (bottom strand, SEQ.ID:6)) containing a HindIII site and codons encoding the Ser, Leu, Asp, Lys and Arg from the C-terminal end of the α-factor pro-peptide was constructed. The α-factor adaptor was ligated to the synthetic HV-1 gene such that the recombinant gene encoded an in-frame α-factor pro-peptide fusion to hirudin. The pUC19 HV-1 plasmid DNA of Preparation 1 was first cleaved with BspMI and the overhanging ends were filled using DNA polymerase I Klenow fragment to create a blunt-ended linear DNA fragment. The linearised fragment was separated from

uncut plasmid on a 1% low gelling temperature agarose gel, excised and extracted from the agarose gel matrix, then further treated with HindIII. The fragment was then ligated to the alpha-factor adaptor described above and annealed prior to ligation. The recombinant ligation products were transformed into competent cells of E. coli strain HW87 (Preparation 1). Ampicillin resistant transformants were analysed by preparation of plasmid DNA, digestion with HindIII and BamHI and agarose gel electrophoresis. A correct recombinant plasmid was called pJC80. The α-factor adaptor - hirudin sequence was removed from pJC80 on a ca. 223 bp HindIII-BamHI DNA fragment (SEQ.ID:7). The fragment was purified on a low gelling temperature agarose gel and ligated to HindIII and BamHI treated pSW6 vector DNA. The recombinant ligation products were transformed into competent cells of E. coli strain HW87. Ampicillin resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis with HindIII and BamHI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pJK1 (Figure 3) was identified. This plasmid is the basic vector used for wild-type hirudin HV-1 expression and was used to derive certain other yeast expression vectors as detailed in the remaining preparations and examples.

PREPARATION 3 - Expression of Hirudin Synthetic Gene

Plasmid expression vector pJK1 of Preparataion 2 was transformed into yeast (S. cerevisiae) strain BJ2168 which has the following genotype:prc-1-407, prb1-1122 pep4-3 leu2 trp1 ura3-52 cir+ using the method of

Sherman F. et al (Methods in Yeast Genetics, Cold Spring Harbor Laboratory, (1986)). All yeast media was as described by Sherman et al. Using 2 litre shake flasks, cultures of yeast containing pJK1 were grown in 1 litre batches of 0.67% synthetic complete medium, yeast nitrogen base, with amino acids minus leucine and 1% glucose as a carbon source. After overnight growth at 30°C, the cells were harvested by centrifugation at 3000 rpm for 10 minutes and resuspended in the same synthetic complete medium except that 1% galactose and 0.2% glucose was used as the carbon source. This induces gene expression from the hybrid PGK promoter. Cells were grown in the induction medium for 3 days. After this period, the supernatant was harvested and assayed for hirudin activity as described in Example 2, Section D, below.

EXAMPLE 1 - Construction of a Hirudin-IEGR-Hirudin Fusion Gene and a Vector for its Expression

A factor Xa-cleavable hirudin fusion protein molecule has been engineered in which two full length hirudin molecules are joined by the peptide linker sequence Ile Glu Gly Arg (See SEQ.ID NO:8). The molecule is designed to be activatable by factor Xa cleavage. The strategy for construction of the hirudin-IEGR-hirudin gene is detailed below.

A gene encoding the hirudin-IEGR-hirudin molecule was constructed by oligonucleotide directed mutagenesis and molecular cloning. Mutagenesis was carried out according to the method of Kunkel et al., Methods in Enzymology, 154, 367-382 (1987). Host strains are described below.

E. coli strains

RZ1032 is a derivative of E. coli that lacks two enzymes of DNA metabolism: (a) dUTPase (dut), the lack of which results in a high concentration of intracellular dUTP, and (b) uracil N-glycosylase (ung) which is responsible for removing mis-incorporated uracils from DNA (Kunkel et al., loc. cit.). A suitable alternative strain is CJ236, available from Bio-Rad Laboratories, Watford WD1 8RP, United Kingdom. The principal benefit is that these mutations lead to a higher frequency of mutants in site directed mutagenesis. RZ1032 has the following genotype:

HfrKL16PO/45[lysA961-62), dut1, ung1, thi1, recA, Zbd-279::Tn10, supE44

JM103 is a standard recipient strain for manipulations involving M13 based vectors. The genotype of JM103 is DELTA (lac-pro), thi, supE,strA, endA, sbcB15, hspR4, F' traD36, proAB, lacIq, lacZDELTAM15. A suitable commercially available alternative E. coli strain is E. coli JM109, available from Northumbria Biologicals Ltd.

Mutagenesis

Prior to mutagenesis it was neccesary to juxtapose two adjacent hirudin genes in an M13 mutagenesis vector. This was accomplished as described below. pJK1 vector DNA of Preparation 2 was prepared and an aliquot treated with restriction endonucleases BglII and BamHI, a ca. 466 bp BglII-BamHI DNA fragment from

this digestion was gel purified and ligated to BamHI treated and phosphatased pJC80 vector DNA of Preparation 2. The recombinant ligation products were transformed into competent cells of E. coli strain HW87 (Preparation 1). Ampicillin (100 μg/ml) resistant clones were analysed by plasmid DNA preparation, restriction endonuclease digestion and gel electrophoresis. Clones with inserts in the desired orientation were identified after digestion with KpnI which released a DNA fragment of ca. 465bp in length. (The products of KpnI digestion were analysed on an agarose gel.) One of the correct clones, pJK002, was used for the remaining constructions, this vector contains a ca. 465 bp KpnI DNA fragment which encodes a C-terminal portion of a first hirudin gene, a complete α-factor pre-pro-peptide sequence and the N-terminal portion of a second hirudin gene. In order to delete the α-factor pre-pro-peptide sequence and to insert DNA encoding a factor Xa-cleavable amino acid linker sequence (IEGR), the ca. 465 bp KpnI DNA fragment was transferred into a bacteriophage mutagenesis vector M13mp18. Plasmid DNA of pJK002 was prepared and a portion was digested with KpnI. The ca. 465 bp KpnI DNA fragment from pJK002 was gel purified and ligated to KpnI treated and phosphatased M13mp18. The recombinant ligation products were transfected into competent cells of E. coli strain JM103. Single stranded DNAs from putative recombinant phage plaques were prepared and analysed by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ. ID NO: 10; see below). A clone pGC609 containing the KpnI fragment in the correct orientation was identified.

The α-factor pre-pro-peptide sequence between the two hirudin sequences of pGC609 was deleted and the DNA encoding the Factor Xa-cleavable amino acid linker (IEGR) inserted by site directed mutagenesis. Single stranded DNA of pGC609 was prepared from E. coli strain RZ1032 and was used as a template for mutagenesis with a 46mer oligonucleotide BB2988: (5'-CAGTCGGTGTAAACAACTCTTCCTTCGATCTGCAGATATTCTTCTG-3') (SEQ. ID NO:9). Single stranded DNAs were prepared from putative mutant plaques and were analysed by dideoxy DNA sequence analysis using an M13 universal sequencing primer (United States Biochemical Corporation. P.O. Box 22400, Cleveland, Ohio 44122. USA. Product No. 70763 5'-GTTTTCCCAGTCACGAC-3'), (SEQ. ID NO:10). A correct clone, pGC610, was identified. To construct the full length hirudin-IEGR-hirudin gene the central core of the fusion molecule encoded on the ca. 210 bp KpnI fragment of pGC610 was cloned into the KpnI site of pJC80 of Preparation 2. Replicative form DNA of pGC610 was prepared and digested with KpnI. The ca. 210 bp KpnI DNA fragment encoding the central core of the hirudin-IEGR-hirudin protein was gel purified and ligated to KpnI treated and phosphatased pJC80 of Preparation 2. The recombinant ligation products were transformed into competent cells of E. coli strain HW87 (Preparation 1). Ampicillin (100 μg/ml) resistant transformants were analysed by preparation of plasmid DNA, restriction endonuclease digestion with PstI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pDB1 was identified as containing a ca. 210 bp DNA fragment after PstI digestion.

To create a vector for the expression of the factor Xa-cleavable hirudin-IEGR-hirudin fusion protein the gene was cloned into the yeast expression vector pSW6 of Preparation 2. Plasmid DNA of pDB1 was treated with HindIII and BamHI and the ca. 420 bp HindIII-BamHI DNA fragment containing the factor Xa-cleavable hirudin-IEGR-hirudin gene was gel purified and ligated to HindIII and BamHI treated pSW6 DNA of Preparation 2. The recombinant ligation products were transformed into competent cells of E. coli strain HW87. Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis with HindIII and BamHI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pDB2 was identified. pDB2 contained the hirudin-IEGR-hirudin gene fused in frame to the yeast α-factor pre-pro-peptide sequence. pDB2 plasmid DNA was prepared and used to transform yeast strain BJ2168 (Preparation 3) according to the method of Sherman F. et al (Methods in Yeast Genetics, Cold Spring Harbor Laboratory, New York (1986)).

EXAMPLE 2 - Purification of Hirudin and Hirudin-IEGR-Hirudin

The procedure of Preparation 3 was generally followed for the expression of hirudin and hirudin-IEGR-hirudin proteins. Hirudin and hirudin-IEGR-hirudin are purified from yeast culture broth. Cells were first removed by centrifugation at 3000 rpm for 10 minutes. The supernatant was then assayed for biological activity using a chromogenic assay (see below, section D). Production levels from shake flask cultures

were routinely between 10-15 mg/litre of culture. The hirudin protein was purified by preparative HPLC (DYNAMAX (Trade Mark) C18, 300 angstroms). The column was first equilibrated in 15% acetonitrile, 0.1% trifluoro acetic acid. Then 2.5-3 mg of hirudin activity as determined by chromogenic assay (section D) was loaded onto the column. The protein was eluted using a 15-40% acetonitrile gradient at 3 ml/minute over 25 min. The purity of the isolated protein was assessed by analytical HPLC (VYDAC (Trade Mark) C18 reverse phase), N-terminal sequence analysis and mono Q FPLC as described below.

A. Assessing Purity by Analytical HPLC

Samples were analysed on a VYDAC (Trade Mark) C18 column (15 x 0.46cm, particle size 5 micron) equilibrated with 10% acetonitrile, 0.1% trifluroacetic acid (TFA). Purified protein (20 μg) was loaded in 10% acetonitrile, 0.1% TFA. Protein was eluted at a flow rate of 1ml/minute using an acetonitrile gradient from 10-40% in 0.1% TFA over 30 minutes. The eluted protein sample was monitored by absorbance at 280 nm.

B. Analysis of Purity by Mono Q FPLC

Samples were analysed on a Mono Q FPLC column (5 x 0.5cm, Pharmacia) equilibrated in 20 mM Tris.HCl pH 7.5. Approximately 15 μg of lyophilised protein was reconstituted in 1ml 20mM Tris.HCl pH 7.5 and loaded onto the column. Protein was eluted using a gradient of 0-250mM NaCl in 20 mM Tris.HCl buffer (pH 7.5) at a flow rate of 1ml/minute over 30 minutes.

C. N-terminal Sequence Analysis

N-terminal sequence analysis was performed by automated Edman degradation using an Applied Biosystems Protein Sequencer, model 471 A (Applied Biosystems, Foster City, California).

Purified material that was greater than 95% pure, was dried down in a SPEEDIVAC (trade mark of Savant Instruments Inc. Hicksville, N.Y. U.S.A.) and reconstituted in 0.5 ml of 0.9% (w/v) saline for assay.

D. Hirudin Anti-thrombin Chromogenic Activity Assay

The ability of hirudin and molecules containing hirudin to inhibit the thrombin catalysed hydrolysis of the chromogenic substrate tosyl-Gly-Pro-Arg-p-nitroanilide (CHROMOZYM TH (trade mark of Boehringer-Mannheim)) was used as an assay to determine their anti-thrombin activity. Protein samples (50 µl) diluted in 0.1M Tris.HCl pH8.5, 0.15 M NaCl, 0.1% (w/v) PEG 6000 were mixed with 50 µl human thrombin (Sigma, 0.8 U/ml in the above buffer) and 50 µl CHROMOZYM TH (2.5mM in water) in 96 well plates (Costar). The plates were incubated at room temperature for 30 minutes. The reaction was terminated by adding 50 µl 0.5 M acetic acid and the absorbance read at 405 nm using an automatic plate reader (Dynatech). Quantitation was performed by comparison with a standard hirudin preparation (recombinant [Lys-47]-HV-2 purchased from Sigma: Sigma Chemical Co. Ltd, Fancy Road, Poole, Dorset BH11 7TG, United Kingdom).

EXAMPLE 3 - Cleavage and Activation of Hirudin-IEGR-Hirudin Fusion Protein

Purified hirudin-IEGR-hirudin fusion protein was incubated with Factor Xa. The reaction was performed at 37°C in a total volume of 150 µl of 0.1M Tris.HCl buffer pH 7.8 and contained 2.06 nmol fusion protein and 0.4 nmol Factor Xa. Analysis of the reaction mixture by sodium dodecyl sulphate-polyacrylamide gel electro- phoresis (SDS-PAGE) demonstrated cleavage to products of a similar size to native hirudin. Reverse phase HPLC analysis of the cleavage reaction as in Example 2, section A, demonstrated the appearance of two new species with retention times (RT) of 17 and 20 minutes compared to 22 minutes for the intact fusion protein.

Measurements of specific activity were made on the products isolated from a cleavage reaction. Using a chromogenic assay according to the method of Example 2, section D, to measure hirudin activity in anti-thrombin units and A 280 nm to determine protein concentration, the following results were obtained: product RT 17 min., 6125 U/mg; product RT 20 min., 5226 U/mg; intact hirudin-IEGR-hirudin, RT 22 min., 2588 U/mg. Cleavage therefore produces an approximate 2-fold increase in specific activity, with the products displaying similar values to that recorded for a recombinant hirudin sample (6600 U/mg) as measured according to the method of Example 2, section D.

Purified cleavage products and the intact fusion protein were subjected to N-terminal sequence analysis.

In each case the sequence obtained was identical to that of native hirudin (HV1), (VVYTD).

It has thus been demonstrated that the hirudin-IEGR-hirudin fusion protein can be cleaved by Factor Xa to produce two products with hirudin activated. Cleavage of the fusion protein is accompanied by activation as the products of cleavage have approximately double the specific activity of the fusion protein.

PREPARATION 4 - Isolation of a streptokinase gene

Streptokinase is secreted by Lancefield's Group C streptococci and cloning of the streptokinase gene from Streptococcus equisimilis strain H46A has been described (Malke,H. and J.J. Ferretti, P.N.A.S. **81** 3557-3561 (1984)). The nucleotide sequence of the cloned gene has been determined (Malke, H., Roe, B. and J.J. Ferretti, Gene **34** 357-362 (1985)). A gene encoding streptokinase has been cloned from S. equisimilis (ATCC 9542 or ATCC 10009) for use in the current invention. Methods that can be used to isolate genes are well documented and the procedure used to isolate the streptokinase gene is summarized in the following protocol.

1. DNA was prepared either from Streptococcus equisimilis (Lancefield's Group C) ATCC 10009 or from ATCC 9542 grown in brain-heart infusion medium (Difco-Bacto Laboratories, PO Box 14B, Central Avenue, E. Mosely, Surrey KT8 OSE, England) as standing cultures. Chromosomal DNA was isolated from

approximately 1.5 ml of cells at a density of $1x10^{11}$ cells/ml. The cells were harvested and washed in 1ml buffer (0.1M potassium phosphate pH 6.2). The pellet was resuspended in 400 µl of the same buffer and 500 units of mutanolysin (Sigma Chemical Company Ltd, Fancy Road, Poole, Dorset BH17 7TG, UK) in 100µl volume was added. This mix was incubated at 37°C for 1 hour. The cells were harvested by centrifugation and again washed in buffer. The cells were resuspended in 500µl of a solution containing 50mM glucose, 10mM EDTA and 25mM Tris HCl pH 8.0 and incubated at 37°C for approximately 1 hour with the mix being shaken gently to prevent the cells settling. A 500µl aliquot of a solution containing 0.4% SDS and proteinase K (100µg/ml) (Sigma Chemical Company Ltd) was added and the mix was incubated at 37°C for 1 hour until it became viscous and clear. The mix was then extracted three times with phenol equilibrated with TE buffer (10mM Tris HCl, 1mM EDTA pH 8.0). The aqueous phase was removed into an eppendorf tube, sodium acetate added to a final concentration of 0.3M and 2.5 volumes of ethanol added. The mix was incubated at -70°C for 1 hour to precipitate the DNA. The DNA was pelleted by centrifugation, washed with 70% ethanol and then resuspended in 200 µl TE buffer.

2. The Polymerase Chain Reaction (PCR) was used to amplify the streptokinase sequence (Saiki R. et al Science, 239, 487-491 (1988)). Two primers were designed based on the published streptokinase sequences. The primer encoding the antisense strand at the 3' end of the gene was a 40mer BB1888 (5'GTTCATGGATCCTTATTTGTCGTTAGGGTTATCAGGTATA 3'), (SEQ.

ID NO:11) which also encoded a BamHI site. The primer encoding the sense strand at the 5' end of the gene encoded an EcoRI site in addition to the streptokinase sequence and was the 40mer BB1887 (5'TCAAGTGAATTCATGAAAAATTACTTATCTTTTGGGATGT 3'), (SEQ ID NO:12). Forty cycles of PCR were performed with the denaturation step at 95°C for 2 minutes, followed by annealing of the primers for 3 minutes at 55°C and extension at 70°C for 4.5 minutes. A sample of the reaction product was analysed on a 0.8% agarose gel. A single amplified DNA fragment at c.a. 1.3 kB, which corresponds to the expected size of the streptokinase gene, was observed.

3. A 30µl sample of the product was digested with the restriction endonucleases EcoRI and BamHI, analysed on a low gelling temperature agarose gel and the c.a. 1.3 kb DNA fragment was isolated from the gel. The band was extracted from the gel and ligated into the plasmid pUC19 which had been cleaved with EcoRI and BamHI to form the plasmid pUC19SK.

The entire ca. 1330 bp EcoRI-BamHI fragment from pUC19SK was sequenced by dideoxy sequence analysis. To facilitate the sequencing, The EcoRI-BamHI DNA fragment of pUC19SK was transferred to M13 sequencing vectors mp18 and mp19 in two halves. A ca. 830 bp EcoRI-HindIII DNA fragment was separately transferred into EcoRI and HindIII treated M13mp18 and M13mp19. The products from these two ligation events were separately transfected into competent cells of E. coli host JM103. Single stranded DNA was prepared and used for dideoxy sequence analysis using the primers listed

in SEQ ID NO: 13 and SEQ ID NO: 10. A ca. 490 bp HindIII-BamHI fragment was gel purified after treatment of pUC19SK with HindIII and BamHI. This DNA fragment was separately ligated to M13mp18 and M13mp19 which had been treated with HindIII and BamHI. The products of these two ligations was transfected into competent cells of E. coli host JM103. Single stranded DNA was prepared and used for dideoxy sequence analysis with the primers shown in SEQ ID NO:13 and SEQ ID NO: 10. The entire sequence of the EcoRI-BamHI PCR derived DNA fragment is shown in SEQ ID NO:14.

EXAMPLE 4 - Construction of Streptokinase Expression Vectors

A number of alternative streptokinase expression vectors have been constructed for expression in either yeast S. cerevisiae or E. coli K12.

1) Vectors for secretion to the periplasm of E. coli K12

Two vectors were designed to enable the secretion of streptokinase to the periplasmic space after expression in E. coli K12. Secretion of streptokinase is desirable to facilitate production of protein with an authentic N-terminus, to ease purification, to reduce potential toxic effects and to limit intracellular proteolysis. Secretion of streptokinase through the E. coli cytoplasmic cell membrane was directed by either the streptokinase signal peptide or the E. coli major outer membrane protein A (OmpA) signal peptide (OmpAL).

A. Secretion using the streptokinase leader

The streptokinase gene of Preparation 4 was transferred into the E. coli expression vector pGC517 (Figure 4). pGC517 contains the regulatable ptac promoter, a ribosome binding site and a synthetic transcriptional terminator. pGC517 was deposited in E. coli K12 at The National Collection of Industrial and Marine Bacteria Limited, 23 St. Machar Drive, Aberdeen, AB2 1RY, Scotland, United Kingdom on 5th December 1990 under Accession No. NCIMB 40343. Genes can be cloned into the expression site of pGC517 on NdeI-BamHI DNA fragments. It was necessary to engineer a NdeI site into the 5′ end of the streptokinase gene to enable subsequent cloning into pGC517. The NdeI site was introduced by site-directed mutagenesis. To construct the vector for the site directed mutagenesis, plasmid DNA of vector pUC19SK of Preparation 4 was prepared and digested with EcoRI and BamHI and the ca. 1.3 Kb EcoRI-BamHI DNA fragment was gel purified and ligated to M13mp18 treated with EcoRI and BamHI. Recombinant ligation products were transfected into competent cells of E. coli strain JM103 (Example 1). Single stranded DNA was prepared from the putative recombinant plaques and analysed by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO: 10 of Example 1). One of the correct recombinant phages was called pGC611. Single stranded DNA of phage pGC611 was prepared from E. coli strain RZ1032 (Example 1) and used as a template for mutagenesis. An NdeI restriction site was introduced by site-directed mutagenesis at the 5′ end of the streptokinase gene such that the NdeI site

overlapped the streptokinase initiation codon. The mutagenesis was performed using a 26-mer BB2175 (5'-GATAAGTAATTTTTCATATGAATTCG-3'), (SEQ ID NO:15). Single stranded DNAs were prepared from putative mutant plaques and were screened by dideoxy sequence analysis using the 18mer sequencing primer BB2358 (5'-CATGAGCAGGTCGTGATG-3'), (SEQ ID NO:16) and a correct clone pGC612 was identified.

To construct an expression vector, the streptokinase gene carrying the newly introduced NdeI site, was cloned into the pGC517 expression vector. Replicative form DNA was prepared from pGC612 and was digested with NdeI and BamHI and the ca. 1.3 kb NdeI-BamHI DNA fragment was gel purified. This fragment was then ligated to NdeI and BamHI treated pGC517 DNA. The recombinant ligation products were transformed into competent cells of E. coli strain JM103. Ampicillin (100 μg/ml) resistant transformants were analysed by plasmid DNA preparation, restriction endonuclease digestion with BglII and BamHI and agarose gel electrophoresis. One of the correct clones, pKJ2, was verified by dideoxy sequence analysis using the sequencing primer BB2358. This vector contains the entire streptokinase gene including the sequences encoding the streptokinase signal peptide leader region and was used for the expression of streptokinase in E. coli.

B. Secretion using the E. coli OmpA leader

As an alternative secretion signal, a DNA sequence encoding the major outer membrane protein A (OmpA)

signal peptide (OmpAL) was fused to the DNA sequence encoding the mature streptokinase protein; see SEQ ID NO:17. A DNA fragment encoding streptokinase was obtained by preparing pUC19SK vector DNA, treating the DNA with EcoRI and filling-in the overhanging single stranded DNA ends with DNA polymerase I Klenow fragment to create a blunt-ended linear DNA fragment. The fragment was next digested with BamHI and the ca. 1.3 kb blunt-ended-BamHI DNA fragment containing the streptokinase gene was gel-purified. The DNA sequence encoding OmpAL is available on an expression vector pSD15. The pSD15 vector contains a gene encoding an insulin like growth factor II gene (IGF-II) fused to the OmpAL signal sequence. pSD15 was deposited in E. coli K12 at The National Collection of Industrial and Marine Bacteria Limited, 23 St. Machar Drive, Aberdeen, AB2 1RY, Scotland, United Kingdom on 5th December 1990 under Accession No. NCIMB 40342. In order to use pSD15 as a vector to provide the OmpAL DNA sequence, pSD15 vector DNA was treated with NheI, the single stranded DNA overhanging ends were filled-in with DNA polymerase I Klenow fragment to create a blunt-ended linear DNA fragment. The linear DNA fragment was next digested with BamHI which removed ca. 123 bp from the 3′ end of the IGF-II gene in pSD15. After restriction endonuclease digestion the cleaved linear DNA fragment was treated with phosphatase, to prevent recircularisation of any partially cut vector DNA and was gel purified then ligated to the blunt-ended-BamHI DNA fragment containing the streptokinase gene. The ligated mixture was transformed into competent cells of E. coli strain HW87 (Preparation 1). Ampicillin (100 μg/ml) resistant

recombinants carrying the streptokinase gene were characterised by preparation of plasmid DNA, restriction endonuclease analysis with BglII and HindIII and agarose gel electropohoresis. A construct of the correct electrophoretic pattern was called pKJ1. Vector pKJ1 contains the DNA encoding OmpAL and streptokinase separated by a region of DNA not required in further constructs. The sequence of the insert DNA in pKJ1 was confirmed by dideoxy sequence analysis with a 44-mer oligonucleotide BB58 (5'-AGCTCGTAGACACTCTGCAGTTCGTTTGTGGTGACCGTGGCTTC-3') SEQ ID NO:18. In order to create a DNA template for the deletion loopout mutagenesis of the unwanted DNA sequence, the BglII to HindIII DNA fragment from pKJ1 was cloned into a vector M13mp19. pKJ1 vector DNA was treated with BglII and HindIII to produce a ca. 1026 bp DNA fragment, which was gel purified and ligated into the polylinker region of M13mp19 replicative form DNA treated with BamHI and HindIII. Ligation products were transfected into competent cells of E. coli strain JM103. Single stranded DNAs were prepared from putative recombinant plaques and a correct clone (pGC600) identified by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO:10, Example 1).

Mutagenesis on template pGC600 was performed using a 30-mer oligonucleotide mutagenesis primer BB2658 (5'-ACCGTAGCGCAGGCCATTGCTGGACCTGAG-3') SEQ ID NO:19. Single stranded DNAs were prepared from putative mutant plaques and a clone, pGC601, containing the required deletion was identified using dideoxy sequence analysis with the M13 universal sequencing

primer (SEQ ID NO: 10). pGC601 contains part of the OmpAL-streptokinase fusion required for the secretion of streptokinase from this signal peptide in E. coli, but DNA encoding the C-terminal portion of streptokinase is absent. In order to reconstruct the streptokinase gene, replicative form DNA from pGC601 was digested with restriction enzymes NdeI and HindIII and the ca. 810 bp NdeI-HindIII DNA fragment containing the DNA sequences encoding OmpAL leader peptide sequence fused to the N-terminal portion of streptokinase was gel purified. pJK2 vector DNA was treated with restriction enzymes NdeI and HindIII followed by treatment with phosphatase and the ca. 3620 bp NdeI-HindIII vector DNA fragment containing the essential vector sequences and the C-terminal portion of the streptokinase gene was gel purified. The ca. 810 bp NdeI-HindIII (pGC601) and ca. 3620 NdeI-HindIII (pKJ2) gel purified DNA fragments were ligated together and the recombinant ligation products were transformed into competent cells of E. coli strain HW1110 (lacIq). The lacIq mutationin this strain enhances repression of transcription from the tac promoter. Any other lacIq strain, for example JM103 could be used instead. The ampicillin resistant transformants were screened by preparation of plasmid DNA followed by restriction endonuclease analysis using NdeI and HindIII. Agarose gel electrophoresis of digestion products was used to identify a correct clone which was called pLGC1. The pLGC1 construct was verified by dideoxy sequence analysis using a 17-mer oligonucleotide BB2753 (5'-GACACCAACCGTATCAT-3'), (SEQ ID NO: 20) to sequence through the BamHI site and primer BB3510 (5'-CACTATCAGTAGCAAAT-3'), (SEQ ID NO:21)

to sequence through the sequence encoding the OmpA leader.

2) Intracellular Expression in E. coli

As streptokinase contains no disulphide bonds there is no requirement for secretion to encourage native protein folding and although streptokinase is naturally secreted, intracellular expression offers several potential advantages such as high yield and inclusion body formation which may facilitate purification. As an alternative production route, an expression vector was designed for intracellular production of streptokinase in E. coli. DNA encoding the amino acids 2 to 21 of the OmpAL signal peptide sequence which was fused to mature streptokinase in pGC601 were deleted by loopout site directed mutagenesis using single stranded DNA of pGC601 with a 31-mer mutagenesis oligonucleotide BB3802 (5'-GAAATACTTACATATGATTGCTGGACCTGAG-3'), (SEQ ID NO:22). In addition to deleting the OmpAL signal peptide coding sequence, BB3802 fused the methionine codon (ATG) of the OmpAL signal peptide sequence to the first codon of mature streptokinase to create the 5'end of gene encoding a Methionyl-streptokinase fusion protein (see SEQ ID NO:23). The ATG codon was used to allow initiation of translation at the correct position. Single stranded DNA was prepared from putative mutant plaques and a clone containing the desired mutation, pGC602 was identified using dideoxy sequence analysis with the M13 universal sequencing primer (SEQ ID NO:10). The C-terminal portion of the

streptokinase gene is missing in pGC602. In order to reconstruct the intact mature streptokinase coding sequence, replicative form DNA from pGC602 was digested with restriction enzymes NdeI and HindIII and the ca. 755 bp NdeI-HindIII DNA fragment encoding the N-terminal portion of the Methionyl-streptokinase protein was gel purified and ligated to the gel purified ca. 3620 bp NdeI-HindIII pLGC2 vector DNA fragment described in Example 6 below. The recombinant ligation mixture was transformed into competent cells of E. coli strain HW1110 (lacIq). Ampicillin (100 μg/ml) resistant transformants were screened by plasmid DNA preparation, restriction endonuclease digestion and agarose gel electrophoresis. A clone , pGC603, with the correct electrophoretic pattern after NdeI and HindIII digestion, was identified. Vector pGC603 was used for the intracellular expression of Methionyl-streptokinase in E. coli strain HW1110.

3) Construction of Expression Vectors for the Secretion of Streptokinase from the Yeast S. cerevisiae

Expression vectors were designed to enable the secretion of streptokinase to the extracellular medium after expression in S. cerevisiae. Secretion of streptokinase is desirable to facilitate production of protein with an authentic N-terminus., to ease purification, to limit intracellular proteolysis and to reduce potential toxic effects on the yeast host. Secretion of streptokinase through the yeast membrane was directed by either the natural streptokinase signal peptide or by fusion of

mature streptokinase to the yeast mating type alpha-factor pre-pro-peptide (a naturally secreted yeast peptide) see SEQ ID NO:24.

A) Secretion of Streptokinase using the Streptokinase Signal Peptide

The streptokinase gene with its natural signal peptide was cloned into the yeast expression vector pSW6 to allow its expression in the yeast S. cerevisiae. Vector DNAs of pKJ2 and pSW6 of Preparation 2 were prepared. Both DNAs were treated with restriction enzymes BglII and BamHI and the ca. 1420 bp DNA fragment from pKJ2 and the ca. 7460 bp vector DNA fragment from pSW6 were gel purified and ligated together. The recombinant ligation products were transformed into competent cells of E. coli strain DH5 (supE44, hsdR17, recA1, endA1, gyrA96, thi-1, relA1), but any other good transforming strain could be used, for example JM109 of Example 1. Ampicillin (100 µg/ml) resistant transformants were analysed by preparation of plasmid DNA, restriction endonuclease digestion with BamHI and HindIII and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pSMD1/111 was used for the expression of streptokinase from its own signal peptide sequence from the yeast S. cerevisiae. Plasmid expression vector pSMD1/111 was transferred into yeast (S. cerevisiae) strain BJ2168 according to the method of Preparation 3.

B) Secretion of Streptokinase using the pre-pro-α-Factor Secretion Leader

A gene fusion to enable the streptokinase gene of Preparation 4 to be expressed in yeast and to be secreted by the yeast mating type α-factor pre-pro-peptide was designed and constructed using site-directed mutagenesis and molecular cloning see SEQ ID NO:24. The construction involved mutagenesis to create an α-factor-streptokinase fusion gene and molecular cloning to reconstruct the DNA sequences encoding the mature streptokinase protein sequence. Single stranded DNA of pGC600 prepared from E. coli strain RZ1032 (Example 1) was used as a mutagenesis template with the 36-mer oligonucleotide BB3624 (5'-GTCCAAGCTAAGCTTGGATAAAAGAATTGCTGGACC-3') SEQ ID NO:25. Single stranded DNA from putative mutant plaques were analysed by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO:10) and a mutant clone, pGC614, with the desired sequence was identified. In pGC614 the OmpA-IGFII-Streptokinase signal peptide encoding sequences of pGC600 have been deleted and the α-factor linker encoding the C-terminal 5 amino acids of the α-factor pro-peptide described in Preparation 2 have been inserted. To reconstruct the streptokinase gene in a yeast expression vector, two stages of genetic manipulation were required. First the C-terminal portion of streptokinase was cloned into a yeast expression vector and this new construct was used to clone in the N-terminal α-factor-streptokinase fusion portion of the gene, thus reconstructing a

mature streptokinase coding region fused to the α-factor pre-propeptide gene. Vector DNAs of pKJ2 and pSW6 (Preparation 2) were prepared and digested with HindIII and BamHI and the ca. 485 bp. DNA fragment from pKJ2 and the ca. 7750 bp. vector DNA fragment from pSW6 were gel purified and ligated. Recombinant ligation products were transformed into competent cells of E. coli strain DH5. Ampicillin resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease digestion with HindIII and BamHI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pSMD1/119 was isolated. It contains DNA encoding the C-terminal portion of streptokinase cloned into a yeast expression vector. The DNA encoding the N-terminal portion of streptokinase and the alpha-factor adaptor sequence were next cloned into pSMD1/119. Replicative form DNA of pGC614 was prepared and treated with HindIII and ligated to pSMD1/119 vector DNA which had been treated with HindIII and phosphatased. The recombinant ligation products were transformed into competent cells of E. coli strain DH5. Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis with DraI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pSMD1/152 gave DraI digestion products of ca. 4750, 1940, 1520 and 700 bp. in length. pSMD1/152 was used for the expression and secretion of streptokinase using the alpha factor pre-pro-sequence from the yeast S. cerevisiae. Plasmid expression vector pSMD1/152 was transferred into yeast (S. cerevisiae) strain BJ2168 according to the method of Preparation 3.

EXAMPLE 5 - Construction of a Gene Encoding a Core Streptokinase Protein

A gene encoding a truncated methionyl streptokinase molecule (aa 16-383) was designed and constructed by oligonucleotide directed loopout deletions and molecular cloning; see SEQ ID NO:26. DNA encoding the amino acids 2 to 21 of the OmpAL signal sequence, the DNA encoding IGF-II, the DNA encoding the streptokinase signal peptide and the first 15 amino acids of the mature streptokinase protein in pGC600 of Example 4B were deleted by loopout mutagenesis using a 33-mer oligonucleotide BB3862:
5'-GAAATACTTACATATGAGCCAATTAGTTGTTAG-3'; SEQ ID NO:27. Single stranded DNA was prepared from E. coli RZ1032 cells infected with pGC600 and used as the template for mutagenesis with primer BB3862. Single stranded DNA was prepared from putative mutant plaques and a clone pGC604 containing the desired deletion was identified by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO:10, Example 1).

Amino acids 384 to 414 were deleted from streptokinase by loopout mutagenesis using a 28-mer oligonucleotide BB3904:
5'-CCCGGGGATCCTTAGGCTAAATGATAGC-3'; SEQ ID NO:28. The template for the mutagenesis was single stranded DNA of M13JK1 of Example 10 containing the ca. 500 bp HindIII-BamHI DNA fragment encoding the 3' end of the streptokinase gene from pUC19SK of Preparation 4. Single stranded DNA from putative mutant plaques was prepared and a clone pGC605 containing the desired deletion was identified by

dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO:10, Example 1).

The intact core streptokinase molecule was reconstructed from the two mutated halves by a two step ligation incorporating the NdeI-HindIII DNA fragment from pGC604 (containing the DNA encoding the N-terminal portion of the core streptokinase molecule) and the HindIII-BamHI DNA fragment from pGC605 (containing the DNA encoding the C-terminal portion of the core streptokinase molecule) into the vector DNA pLGC2 of Example 6 below. First the pGC604 DNA was digested with NdeI and HindIII. A DNA fragment of ca. 710 bp. was gel purified. Vector DNA was prepared from pLGC2 of Example 6 and treated with NdeI and HindIII and phosphatased. The linear vector DNA was gel purified and the two fragments were ligated together. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110. Ampicillin (100 μg/ml) resistant transformants were screened for the required clone by preparation of plasmid DNA, restriction endonuclease analysis with NdeI and HindIII followed by agarose gel electrophoresis of the digestion products. One construct with the correct electrophoretic pattern, pGC617, was identified.

To clone the DNA encoding the C-terminal portion, the same vector DNA (pLGC2) was treated with HindIII and BamHI and phosphatased. The pGC605 DNA was treated with HindIII and BamHI and a ca. 402 bp DNA fragment was gel purified and ligated into the HindIII and BamHI treated pLGC2 vector DNA. The recombinant ligation

products were transformed into competent cells of E. coli strain HW1110. Ampicillin (100 μg/ml) resistant transformants were screened for the required clone by preparation of plasmid DNA, restriction endonuclease analysis with BamHI and HindIII, and agarose gel electrophoresis of the digestion products. One construct with the correct electrophoretic pattern pGC618 was identified. Finally, to reconstruct the intact core streptokinase gene from the two halves, pGC617 DNA was treated with HindIII and BamHI and the ca. 402 bp HindIII-BamHI fragment from pGC618 ligated to it. pGC618 DNA was digested with HindIII and BamHI and a ca. 402 bp HindIII-BamHI DNA fragment was gel purified. pGC617 vector DNA was also treated with HindIII and BamHI and a ca. 402 bp HindIII-BamHI DNA fragment from pGC618 was ligated into it. The ligation products were transformed into competent cells of E. coli strain HW1110. Ampicillin resistant transformants were screened by preparation of plasmid DNA restriction endonuclease analysis with BamHI and HindIII and agarose gel electrophoresis. A correct construct, pGC606, was identified.

EXAMPLE 6 - Construction of Expression vectors containing a Thrombin Cleavable Streptokinase-Streptokinase Fusion Gene

1) Construction of a Secretion Vector for the Expression of a Thrombin Cleavable Streptokinase-Streptokinase Fusion

A gene encoding an OmpAL streptokinase-streptokinase fusion linked by a thrombin cleavable linker sequence

VELQGVVPRG, identical to that at the thrombin cleavage site in Factor XIII, was designed and constructed by site directed mutagenesis and molecular cloning (SEQ ID NO:29). A ca. 1.3 Kb EcoRI-BamHI DNA fragment containing a streptokinase gene was gel purified after treatment of the pUC19SK vector DNA of Preparation 4 with EcoRI and BamHI. A second DNA fragment encoding a streptokinase gene was gel purified after BglII and SalI digestion of the pKJ1 vector DNA of Example 4. A trimolecular ligation was carried out between these two fragments and EcoRI and SalI treated pGC517 vector DNA described in Example 4, section 1A. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110 (laqIq). Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis with EcoRI and SalI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern (pSD93) was identifed. pSD93 contains two tandem copies of the streptokinase gene separated by a sequence containing the bacteriophage lambda gene cII ribosome binding site, and encoding the OmpA signal peptide sequence, the streptokinase signal peptide sequence and the 5′ part of the IGF-II sequence from pKJ1. To remove this unwanted intervening sequence and to replace it with the desired thrombin cleavable linker sequence a part of pSD93 was transferred into an M13 mutagenesis vector for mutagenesis. Plasmid pSD93 DNA was digested with HindIII and a ca. 1530 bp DNA fragment gel purified and ligated to HindIII treated and phosphatased replicative form M13mp18 DNA. The recombinant ligation products were

transformed into competent cells of E. coli strain JM103 (Example 1). There are two possible fragment orientations in such a construction. The orientation of the clones was determined by preparation of replicative form DNA and analysing the DNA fragments produced after XmnI digestion and agarose gel electrophoresis. One of the clones pSD95 which contained the fragment in an inverted orienation (thus preventing translation readthrough by virtue of fusion to the α-fragment of ß-galactosidase expressed from the M13 mutagenesis vector) was used for mutagenesis. Single stranded DNA template was prepared from pSD95 and used for site directed mutagenesis. The primer used was a 63-mer oligonucleotide BB2938:
(5'-GATAACCCTAACGACAAAGTAGAGCTGCAGGGAGTAGTTCCTCGTGGAAT-TGCTGGACCTGAG-3') (SEQ ID NO:30) designed to loop out the gene cII ribosome binding site, the OmpAL IGF-II sequence, the streptokinase signal peptide sequence in pSD95 and to insert a DNA sequence encoding a thrombin cleavable amino acid sequence. Single stranded DNAs were prepared from putative mutant plaques and a correct mutant pGC607 was identified using dideoxy sequence analysis with primer BB2753 (SEQ ID NO:20) of Example 4. Replicative form DNA of pGC607 was prepared and was digested with HindIII and the ca. 1277 bp HindIII DNA fragment gel purified and ligated to HindIII treated and phosphatased pLGC1 vector DNA of Example 4. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110. Ampicillin resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis using HindIII

and agarose gel electrophoresis. This cloning rebuilds the gene encoding a thrombin cleavable streptokinase-streptokinase fusion in an expression vector. A clone (pLGC2) carrying the insert in the sense orientation was identified by dideoxy sequence analysis using primers BB2754 (5'-GCTATCGGTGACACCAT-3') SEQ ID NO:31 and BB3639 (5'-GCTGCAGGGAGTAGTTC-3') SEQ ID NO:32. pLGC2 was used for the expression of thrombin cleavable streptokinase-streptokinase fusion protein in E. coli HW1110.

2) Construction of a Vector for the Intracellular Expression of a Thrombin Cleavable Streptokinase-Streptokinase Fusion Gene.

A thrombin cleavable methionyl-streptokinase-streptokinase gene was designed and constructed by molecular cloning. The gene was constructed from the methionyl-streptokinase gene of Example 4 and the HindIII DNA fragment from pGC607 of Example 6, encoding the C-terminal portion of a first streptokinase molecule, a thrombin cleavable linker and an N-terminal portion of a second streptokinase molecule.

Replicative form DNA of pGC607 was prepared and was digested with HindIII and the ca. 1277 bp HindIII DNA fragment was gel purified and ligated to HindIII treated and phosphatased pGC603 vector DNA of Example 4. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110 (lacIq). Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid

DNA, restriction endonuclease analysis with HindIII, BamHI and PstI and agarose gel electrophoresis of the digestion products. One construct with the correct electrophoretic pattern pLGC3, was used for the intracellular expression of a thrombin cleavable methionyl-streptokinase-streptokinase fusion protein.

EXAMPLE 7 - Construction of a Thrombin Cleavable Core Streptokinase-core Streptokinase Fusion Gene

A gene encoding a core methionyl-streptokinase-core streptokinase fusion linked by a thrombin cleavable linker sequence VELQGVVPRG, identical to that at the thrombin cleavage site in Factor XIII, was designed and constructed by site directed mutagenesis and molecular cloning see SEQ ID NO:33. The core streptokinase-core streptokinase fusion gene was constructed from the core streptokinase monomer gene of Example 5 and a HindIII DNA fragment containing the C-terminal portion of a core streptokinase gene, a thrombin-cleavable linker and an N-terminal portion of a core streptokinase gene. To construct the HindIII DNA fragment containing the appropriate deletions and encoding a thrombin-cleavable linker, pGC607 of Example 6 was used as a template for oligonucleotide directed mutagenesis. A 61-mer oligonucleotide BB3861:
(5'-GCTATCATTTAGCCGTAGAGCTGCAGGGAGTAGTTCCTCGTGGAAGCCAA-TTAGTTGTTAG-3') SEQ ID NO:34 was used to delete the streptokinase amino acids 384 to 414, to reconstruct the thrombin cleavable linker sequence VELQGVVPRG and to delete the first 15 amino acids of the N-terminus of streptokinase. Single stranded DNA from putative

mutant plaques was prepared and a correct clone, pGC608, was identified by dideoxy sequence analysis using sequencing primer BB2753 of example 8. Replicative form DNA was prepared from pGC608 and used in further construction.

To construct an intact core methionyl-streptokinase-core-streptokinase fusion, pGC608 DNA was treated with HindIII and the ca. 1140 bp HindIII DNA fragment encoding the C-terminal portion of the core streptokinase molecule, the thrombin cleavable linker sequence and the N-terminal portion of a core streptokinase molecule, was gel purified and ligated to the vector DNA of pGC606 of Example 5 after treatment with HindIII and phosphatase. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110 (lacIq). Ampicillin (100 μg/ml) resistant transformants were analysed by zymography as described in Example 11 below. A correct clone pLGC4, was identified.

EXAMPLE 8 - Construction of a Factor Xa-Cleavable Hirudin-IEGR-Streptokinase Fusion Gene

A hirudin-streptokinase fusion has been designed in which a full length hirudin molecule is joined to full length streptokinase via an IEGR linker sequence cleavable by factor Xa; see SEQ ID NO:35. The gene encoding the hirudin-streptokinase protein was constructed by site directed mutagenesis and molecular cloning. In order to juxtapose the hirudin and streptokinase genes, the DNA fragments encoding these genes were ligated together. The streptokinase

gene from plasmid pKJ2 of Example 4 was isolated by gel purification of a ca. 1.4 kbp DNA fragment after digestion of pKJ2 vector DNA with BglII and BamHI. This DNA fragment contains all of the streptokinase gene together with the DNA encoding the streptokinase signal peptide sequence. This DNA fragment was then ligated to BamHI treated pJK1 DNA of Preparation 2 which contains the hirudin encoding DNA sequence. The recombinant ligation products were transformed into competent cells of E. coli strain HW1110 (lacIq). Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease digestion with HindIII and agarose gel electrophoresis. There are two possible orientations for the insert in this cloning event and correct clones were identified as those which released a ca. 1080 bp DNA fragment after HindIII digestion as analysed on agarose gels. One such clone pJK3, which contains the hirudin gene separated from the streptokinase gene by the streptokinase signal peptide sequence, was used in subsequent manipulations. To create a template for mutagenesis to delete the intervening sequences and to insert the DNA encoding the factor Xa cleavable linker sequence, the hirudin-streptokinase portion of pJK3 was transferred to a mutagenesis vector M13mp18. Plasmid DNA of pJK3 was digested with KpnI and BamHI and the ca. 1490 bp DNA fragment gel purified and ligated to KpnI and BamHI treated M13mp18 replicative form DNA. The recombinant ligation products were transfected into competent cells of E. coli JM103 (Example 1). Single stranded DNA was prepared from putative recombinant plaques and a correct clone

pSMD1/100 (1.1) was identified. To delete the streptokinase signal peptide sequence and to insert the DNA encoding the factor Xa linker sequence single stranded DNA of pSMD1/100 (1.1) was used as a template for mutagenesis with a 46-mer oligonucleotide BB3317: (5'-CACTCAGGTCCAGCAATTCTACCTTCGATCTGCAGATATTCTTCTG-3') SEQ ID NO:36. Single stranded DNA from putative mutant plaques were prepared and a mutant pGC615 was identified by DNA sequence analysis using the sequencing primer BB3510 (5'-CACTATCAGTAGCAAAT-3') SEQ ID NO:37. pGC615 contains the C-terminal portion of the hirudin gene linked to the mature streptokinase protein coding sequence. In order to reconstruct the hirudin gene, replicative form DNA of pGC615 was treated with KpnI and BamHI, the ca. 1320 bp DNA fragment gel purified and ligated to KpnI and BamHI treated pJC80 of Preparation 2. The recombinant ligation products were transformed into competent cells of E. coli strain DH5 (Example 4). Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis with KpnI, BamHI and HindIII and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pSMD1/139 was identified. This plasmid contains DNA encoding the complete factor Xa cleavable hirudin-streptokinase fusion molecule.

EXAMPLE 9 - Construction of a Vector for the Expression of a Factor Xa Cleavable Hirudin-IEGR-Streptokinase Fusion Molecule

To construct a vector for the expression of the

hirudin-IEGR-streptokinase gene, DNA of pSMD1/139 of Example 8 was treated with HindIII and a ca. 963 bp DNA fragment encoding part of the yeast alpha factor secretion signal, all of hirudin, the factor Xa linker and the 5' part of streptokinase as far as the internal HindIII site in the streptokinase sequence was gel purified. This fragment was then ligated to HindIII treated and phosphatased DNA of pSMD1/119 of Example 4. The recombinant ligation products were transformed into competent cells of E. coli strain DH5 (Example 4). Ampicillin resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease digestion with KpnI and BamHI and agarose gel electrophoresis. It is possible to obtain two orientations of the HindIII insert and one clone in the correct orientation pSMD1/146 was identified as releasing a ca. 1311 bp fragment after KpnI and BamHI treatment. pSMD1/146 contains the full length fusion gene under the control of the regulatable PAL promoter described in Preparation 2, and has been designed for the regulated expression and secretion of the factor Xa-cleavable hirudin-streptokinase fusion protein. pSMD1/146 plasmid DNA was prepared and used to transform yeast strain BJ2168 (Preparation 3) according to the method of Sherman, F. et al., (Methods in Yeast Genetics, Cold Spring Harbor Laboratory (1986)).

EXAMPLE 10 - Construction of a Factor Xa Cleavable Streptokinase-IEGR-Hirudin Fusion Gene and its Expression Vector

A gene encoding a streptokinase-hirudin fusion protein linked via a Factor Xa cleavage site (IEGR) was constructed by site-directed mutagenesis and molecular cloning SEQ ID NO:38. In order to juxtapose the streptokinase and hirudin genes, DNA fragments encoding these two gene were ligated together. The pUC19SK vector DNA of Preparation 4 was prepared and treated with HindIII and BamHI and the ca. 500 bp DNA fragment containing the 3' end of the streptokinase gene was gel purified. This fragment was ligated to M13mp19 replicative form DNA treated with HindIII and BamHI. The recombinant ligation mixture was transfected into competent cells of E. coli strain JM103 (Example 1). Single stranded DNA was prepared from putative recombinant plaques and the required clone M13JK1 identifed by dideoxy sequence analysis using the M13 universal sequencing primer (SEQ ID NO:10, Example 1). M13JK1 contains the C-terminal portion of the streptokinase gene. The α-factor hirudin gene was then cloned into M13JK1 to juxtapose both sequences. Plasmid DNA of pJK1 of Preparation 2 was digested with BglII and BamHI and a ca. 465bp DNA fragment encoding the α-factor hirudin fusion was gel purified. This DNA fragment was then ligated to BamHI treated replicative form DNA of M13JK1. The recombinant ligation products were transfected into competent cells of E. coli strain JM103. Single stranded DNA from putative recombinant plaques were prepared and a correct clone

SMD1/100.3 identified by dideoxy sequence analysis using M13 universal sequencing primer (SEQ ID NO:10, Example 1. SMD1/100.3 contains the C-terminal portion of the streptokinase gene and the complete hirudin gene separated by the α-factor encoding sequence described in Preparation 2. In order to delete this sequence and replace it with a factor Xa-cleavable linker sequence, SMD1/100.3 was used as a template for site-directed mutagenesis. Single stranded DNA of SMD1/100.3 was prepared and used for mutagenesis using a 47-mer mutagenesis primer BB3318: (5'-TCGGTGTAAACAACTCTTCTACCTTCGATTTTGTCGTTAGGGTTATC-3") (SEQ ID NO:40). Single stranded DNA from putative mutant plaques were prepared and the required mutation pGC616 identified by dideoxy sequence analysis using the sequencing primer BB2018: (5'-GCGGCTTTGGGGTACCTTCACCAGTGACACATTGG-3') (SEQ ID NO:2). pGC616 contains an additional mutation inadvertently introduced by the mutagenesis procedure. This was corrected by a further mutagenic step. Single stranded DNA of pGC616 was prepared and used as a template for mutagenesis with a 21-mer oligonucleotide BB3623 (5'-GTGTAAACAACTCTACCTTCG-3') (SEQ ID NO:40). Single stranded DNA from putative mutant plaques was prepared and a correct clone pGC620 identified by dideoxy sequence analysis with the sequencing primer BB2018 (SEQ ID NO:2). pGC620 contains the C-terminal portion of the streptokinase gene and the complete hirudin gene fused via DNA encoding a factor Xa-cleavable linker. The intact factor Xa-cleavable streptokinase-hirudin fusion gene was reconstructed in two steps. The C-terminal streptokinase-hirudin sequence from pGC620 was cloned

into the yeast expression vector pSW6 of Preparation 2 and then the N-terminal portion of streptokinase was cloned into the new vector to create the full length streptokinase-hirudin fusion gene.

Replicative form DNA of pGC620 was treated with HindIII and BamHI and a ca. 710 bp HindIII-BamHI DNA fragment encoding the 3′ end of streptokinase, the intervening factor Xa-cleavable linker DNA sequence and all of the hirudin gene was gel purified. This ca. 710 bp DNA fragment was ligated to pSW6 of Preparation 2 digested with HindIII and BamHI. The recombinant ligation products were transformed into competent cells of E. coli strain DH5 (Example 4). Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease analysis using HindIII and BamHI and agarose gel electrophoresis. A clone with the correct electrophoretic pattern pSMD1/143 was identified. The intact fusion gene was then constructed by cloning the N-terminal portion of α-factor-streptokinase into pSMD1/143. Replicative form DNA of pGC614 of Example 4 was treated with HindIII and the ca. 750 bp DNA fragment containing the N-terminal portion of α-factor-streptokinase gel purified and ligated to HindIII treated and phosphatased pSMD1/143 vector DNA. The recombinant ligation products were transformed into competent cells of E. coli strain DH5. Ampicillin (100 μg/ml) resistant transformants were screened by preparation of plasmid DNA, restriction endonuclease digestion with DraI and agarose gel electrophoresis. A clone in the

correct orientation pSMD1/159 was identified as giving rise to 4 fragments of sizes of about 4750 bp, 2140 bp, 1526 bp, and 692 bp after DraI digestion. pSMD1/159 was used for the expression of the factor Xa-cleavable streptokinase-hirudin fusion protein. pSMD1/159 plasmid DNA was prepared and used to transform yeast strain BJ2168 (Preparation 5) according to the method of Sherman, F. et al., (Methods in Yeast Genetics, Cold Spring Harbour Laboratory (1986)).

EXAMPLE 11 - Expression of Monomer Streptokinase Constructs

Expression

Competent cells of E. coli strain JM103 (Example 1) were transformed with DNA of the streptokinase expression vectors of Examples 4, 5, 6 and 7. The lacIq gene in the expression host is desirable to repress transcription from the tac promoter used in all of the E. coli expression constructs. All media for the growth of recombinant E. coli strains were as described in Maniatis et al. Using 1 litre shake flasks, cultures of recombinant E. coli containing streptokinase expression vectors were grown in 250 ml batches of 2TY medium containing 100 μg/ml of carbenicillin at 37°C in an orbital shaker. The optical density of the cultures were monitored at 600 nm. When the culture reached an OD 600 nm of 0.5, expression from the tac promoter was induced by the addition of isopropyl-ß-D-thiogalactoside (IPTG) to a final concentration of 1 mM. After growth for 30 to 240 min the cells were harvested by centrifugation.

SDS-PAGE Separation

The ability of the recombinant E. coli cells to express streptokinase was assayed using zymography. The quantity and molecular weight of streptokinase activity of an E. coli culture was estimated by the following protocol. A 1 ml aliquot of the culture was removed, the cells were harvested by centrifugation (14 000xg) for 5 mins and resuspended in 200 µl of loading buffer (125 mM Tris.HCl pH 6.8, 10% glycerol (w/v), 0.01% (w/v) bromophenol blue, 1% (v/v) 2-mercaptoethanol, 6M urea). An aliquot of this mixture was applied to an SDS-PAGE gel and the protein separated by electrophoresis. The quantity of protein loaded onto the gel was varied by altering the size of the aliquot according to the optical density of the culture upon harvesting. Generally, 10 µl of the mixture from a culture of OD 600 nm of 1.0 was used for each lane.

Zymography

After electrophoresis the polyacrylamide gel was washed in 2% (w/v) Triton X-100 (3x20 mins) followed by water washes (3x20 mins) to remove the SDS and allow renaturation of the streptokinase molecule.

The washed SDS-PAGE gel was then overlayed with an agarose mixture prepared as follows. 200 mg of agarose was dissolved in 18 ml distilled and deionised water (dH_2O) and allowed to cool to 55-60°C. To this 200 mg of MARVEL (trade mark of Premier Brands, U.K. Ltd. P.O. Box 171, Birmingham, B30 2NA, U.K.) (casein) dissolved

in 2 ml of dH_2O, 1 ml of 1M Tris.HCl pH 8.0 and 600 µl of 5M NaCl were added. Just before pouring over the washed SDS-PAGE gel, 700 µl of plasminogen at 300 µg/ml (Sigma P-7397 10 mg/ml in 50 mM Tris.HCl pH 7.5) was added and mixed thoroughly. The mixture was poured over the gel and once set was incubated at 37°C for 2 hours when it could be inspected. Plasminogen activator activity (streptokinase activity) was detected by plasmin digestion of the opaque casein containing overlay which produced clear zones. The position of the zones on the gel was directly related to the size of the active molecules.

The recombinant E. coli JM103 strains containing monomer streptokinase expression vectors pKJ2 of Example 4 and pLGC1 of Example 4 both expressed streptokinase activity with a molecular weight of approximately 47 kDa (Figure 5).

EXAMPLE 12 - Expression of a Thrombin Cleavable Streptokinase-Streptokinase Fusion Protein.

A recombinant E. coli HW1110 (lacIq) strain (Example 1) containing pLGC2 of Example 6, the thrombin cleavable streptokinase- streptokinase fusion gene, was expressed and analysed according to the expression and zymography protocols of Example 11. The E. coli JM103/pLGC2 strain expressed streptokinase activities of several molecular weights, predominantly of 110 kDa and 47 kDa (Figure 5). Cleavage analysis is described in Example 13 below.

EXAMPLE 13 - Cleavage of the Thrombin Cleavable Streptokinase-streptokinase Fusion Protein by Thrombin

Using 1 litre shake flasks, a 3 litre culture of E. coli JM103 (Example 1) containing pLGC2 of Example 6 was grown in 500 ml batches in 2TY medium containing 100 mcg/ml carbenicillin at 37°C with vigorous shaking in an orbital shaker. When the optical density of the cultures reached an O.D. 600 nm of 0.5 the expression of the streptokinase-streptokinase fusion protein was induced by the addition of IPTG to a final concentration of 1 mM. The cultures were incubated at 37°C with vigorous shaking for a further 4 hours when they were harvested by centrifugation at 8,000 r.p.m. for 10 mins. The cells were resuspended in 10 ml of ice cold TS buffer (10 mM Tris.HCl pH 7.5, 20% (w/v) sucrose). 348 µl of 0.5 M EDTA was added and the mixture incubated on ice for 10 mins. The cells were harvested by centrifugation at 8,000 r.p.m. for 5 min at 4°C and the supernatant discarded. The cells were resuspended in 6.25 ml of ice cold sterile H_2O and incubated on ice for 10 min. The cells were harvested by centrifugation at 8,000 rpm. for 5 min at 15,000 g for 30 min at 4°C and the supernatant discarded. The cells were resuspended in 48 ml of ARG buffer (20 mM Tris.HCl pH 7.5, 10 mM $MgCl_2$, 10mM 2-b-mercaptoethanol, 0.5 mM phenylmethyl sulphonyl fluoride, .12 mcM N-tosyl-L-phenylalanine chloromethyl ketone) and sonicated on ice (6 x 30 sec. bursts on maximum power, MSE SONIPREP 150 (trade mark)). The cell sonicate was centrifuged at 15,000 g for 30 min at 4°C. The supernatant was decanted and assayed for streptokinase

activity using the S2251 (KabiVitrum Ltd, KabiVitrum House, Riverside Way, Uxbridge, Middlesex, UB8 2YF, UK) chromogenic assay for the streptokinase activation of plasminogen. S2251 is a specific tripeptide chromogenic substrate for plasmin. 25 µl of 0.1 M Tris.HCl pH 8.0 was placed in wells 2 to 12 of 96 well plates. Aliquots of the sample (25 µl) were placed in wells 1 and 2, and two-fold dilutions made by mixing and pipetting from wells 2 to 3, 3 to 4 and so on to well 11. A 100 µl aliquot of a plasminogen/S2251 mixture (40 µl plasminogen 300 µg/ml, 220 µl S2251 1 mg/ml, 1.04 ml 0.1 M Tris.HCl pH 7.5) was added to each well and the plate incubated at 37°C for 30 min. The reaction was terminated by the addition of 50 mcl of 0.5 M acetic acid. The absorbance was read at 405 nM using an automatic plate reader. Quantification was performed by comparison with a standard streptokinase preparation. The analysis showed that the supernatant contained approximately 2560 u/ml of streptokinase activity.

Solid ammonium sulphate was slowly added to the supernatant to 15% saturation (4.03 g) and allowed to dissolve for 15 min at room temperature. The mixture was then centrifuged for 30 min at 15,000 g at room temperature. The supernatant was decanted and additional solid ammonium sulphate was added to 40% saturation (7.27 g), and allowed to dissolve. The mixture was centrifuged for 30 min at 15,000 g at room temperature and the supernatant discarded. The pelleted protein (the 15-40% cut), was resuspended in 10 ml of ARG buffer. A portion of the 15-40% cut was assayed using the S2251 chromgenic assay and was found to contain 18,432 u/ml of streptokinase activity.

The ability of thrombin to cleave the streptokinase-streptokinase fusion protein at the thrombin cleavable linker was assessed by an in vitro cleavage assay and zymography. A 5 µl aliquot of the 15-40% cut was mixed with 45 µl of ARG buffer to dilute the protein ten-fold. 10 µl of this mixture was incubated with 5 u/ml of thrombin in a final volume of 50 µl at 37°C for 14 hours. Aliquots (10 µl) of the thrombin cleavage reactions were analysed by zymography according to the method of Example 11. The results are shown in Figure 6. The streptokinase-streptokinase fusion protein (Mr 110 kDa), was quantitatively cleaved whilst the lower molecular weight streptokinase activity (Mr 47 kDa) was not cleaved by thrombin. Thus the streptokinase-streptokinase fusion protein is cleavable by thrombin.

EXAMPLE 14 - Expression of a Factor Xa Cleavable Streptokinase-IEGR-hirudin Fusion Gene

Plasmid expression vector pSMD1/159 of Example 10 was transferred into yeast (S. cerevisiae) strain BJ2168 according to the method of Preparation 3. Using 500 ml shake flasks, cultures of yeast containing pSMD1/159 were grown in 100 ml batches of 0.67% synthetic complete medium yeast nitrogen base, with amino acids minus leucine and 1% glucose as a carbon source. After overnight growth at 30°C, the cells were harvested by centrifugation at 3,000 rpm for 10 min and resuspended in the same synthetic complete medium except having 1% galactose and 0.2% glucose as the carbon source and the addition of sodium phosphate

(to 50 mM) pH 7.2. This induces the expression of the streptokinase-hirudin fusion gene from the hybrid PGK promoter. Cells were grown in the induction medium for 3 days. After this period, the supernatant was harvested by centrifugation. The broth was assayed for both streptokinase activity according to the S2251 assay procedure of Example 13 and hirudin activity according to the thrombin inhibition assay of Example 2. Both activities were detected and secreted to the medium.

EXAMPLE 15 - Expression of a Factor Xa Cleavable Hirudin-IEGR-Streptokinase Fusion Gene

Plasmid expression vector pSMD1/146 of Example 9 was transferred into yeast (S. cerevisiae) strain BJ2168 according to the method of Preparation 3. The culture was incubated, expressed, harvested and the hirudin and streptokinase activities assayed according to the methods of Examples 2 and 13. Both streptokinase and hirudin activities were detected and secreted to the medium.

SEQUENCE LISTINGS

SEQ.ID NO:1

SEQUENCE TYPE: nucleotide with corresponding protein
SEQUENCE LENGTH: 201 base pairs
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
SOURCE: synthetic
FEATURES: hirudin type HV-1 gene
from 195 to 201 bp non-translated stop codons

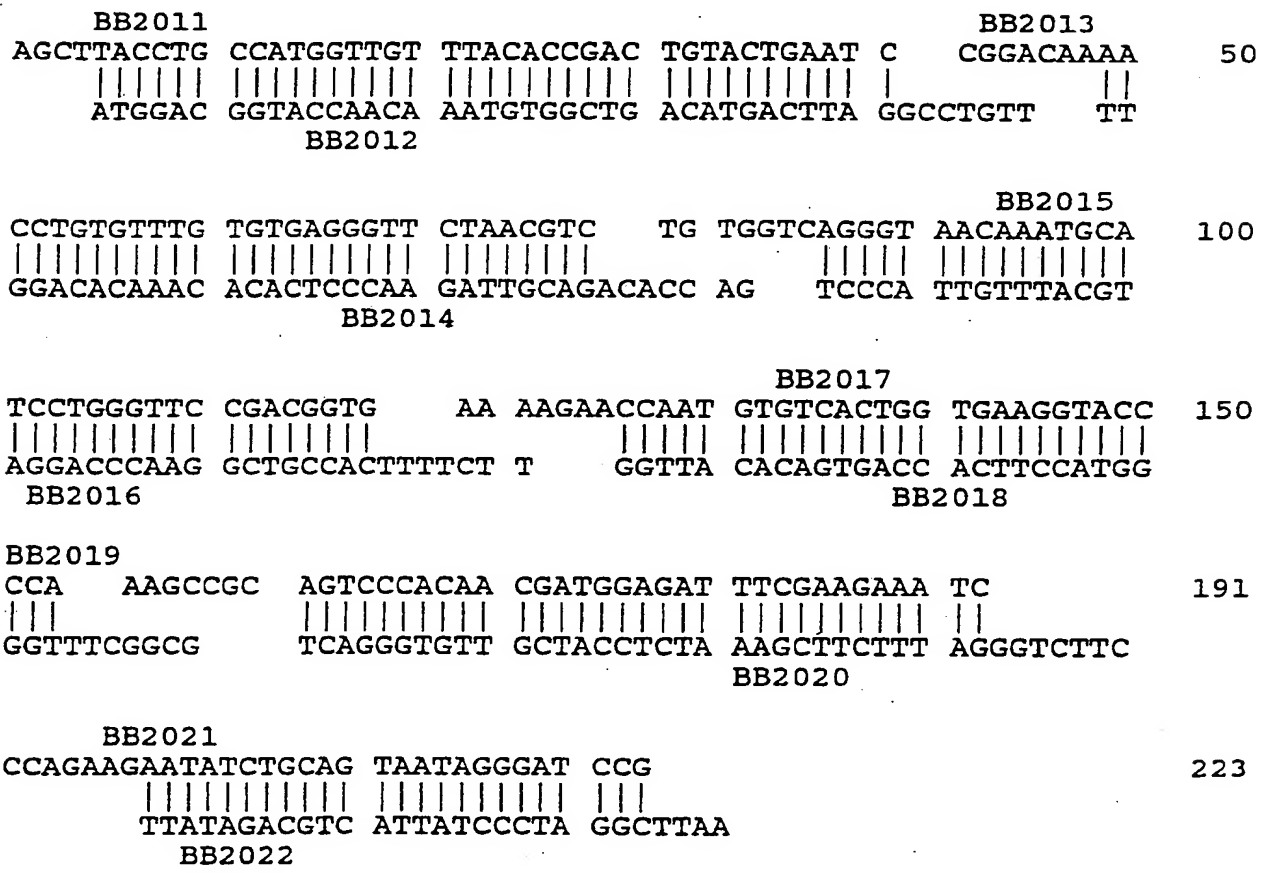
SEQUENCE:

```
GTT GTT TAC ACC GAC TGT ACT GAA TCC GGA CAA AAC CTG TGT TTG 45
CAA CAA ATG TGG CTG ACA TGA CTT AGG CCT GTT TTG GAC ACA AAC
Val Val Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Leu Cys Leu
                  5                  10                  15

TGT GAG GGT TCT AAC GTC TGT GGT CAG GGT AAC AAA TGC ATC CTG  90
ACA CTC CCA AGA TTG CAG ACA CCA GTC CCA TTG TTT ACG TAG GAC
Cys Glu Gly Ser Asn Val Cys Gly Gln Gly Asn Lys Cys Ile Leu
             20                  25                  30

GGT TCC GAC GGT GAA AAG AAC CAA TGT GTC ACT GGT GAA GGT ACC 135
CCA AGG CTG CCA CTT TTC TTG GTT ACA CAG TGA CCA CTT CCA TGG
Gly Ser Asp Gly Glu Lys Asn Gln Cys Val Thr Gly Glu Gly Thr
         35                  40                  45

CCA AAG CCG CAG TCC CAC AAC GAT GGA GAT TTC GAA GAA ATC CCA 180
GGT TTC GGC GTC AGG GTG TTG CTA CCT CTA AAG CTT CTT TAG GGT
Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe Glu Glu Ile Pro
     50                  55                  60

GAA GAA TAT CTG CAG TAATAG                                  201
CTT CTT ATA GAC GTC ATTATC
Glu Glu Tyr Leu Gln
         65
```

**** END OF SEQ ID NO: 1 *****

SEQ. ID NO:2

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 223 base pairs
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: synthetic DNA
SOURCE: synthetic
FEATURES: oligomers designed for construction of synthetic type HV-1 gene.
SEQUENCE:



**** END OF SEQ ID NO: 2 *****

SEQ. ID NO:3

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 19 base pairs
FEATURES: Universal sequencing primer complementary to the universal primer region of pUC19.

SEQUENCE:

CAGGGTTTTC CCAGTCACG 19

**** END OF SEQ ID NO: 3 *****

SEQ ID NO:4

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 7859 base pairs
STRANDEDNESS: single
TOPOLOGY: circular
SOURCE: experimental
FEATURES: Sequence of plasmid pSW6
SEQUENCE:

TTCCCATGTC TCTACTGGTG GTGGTGCTTC TTTGGAATTA TTGGAAGGTA 50
AGGAATTGCC AGGTGTTGCT TTCTTATCCG AAAAGAAATA AATTGAATTG 100
AATTGAAATC GATAGATCAA TTTTTTTCTT TTCTCTTTCC CCATCCTTTA 150
CGCTAAAATA ATAGTTTATT TTATTTTTTG AATATTTTTT ATTTATATAC 200
GTATATATAG ACTATTATTT ACTTTTAATA GATTATTAAG ATTTTTATTA 250
AAAAAAAATT CGTCCCTCTT TTTAATGCCT TTTATGCAGT TTTTTTTTCC 300
CATTCGATAT TTCTATGTTC GGGTTTCAGC GTATTTTAAG TTTAATAACT 350
CGAAAATTCT GCGTTTCGAA AAAGCTCTGC ATTAATGAAT CGGCCAACGC 400
GCGGGGAGAG GCGGTTTGCG TATTGGGCGC TCTTCCGCTT CCTCGCTCAC 450
TGACTCGCTG CGCTCGGTCG TTCGGCTGCG GCGAGCGGTA TCAGCTCACT 500
CAAAGGCGGT AATACGGTTA TCCACAGAAT CAGGGGATAA CGCAGGAAAG 550
AACATGTGAG CAAAAGGCCA GCAAAAGGCC AGGAACCGTA AAAAGGCCGC 600
GTTGCTGGCG TTTTTCCATA GGCTCCGCCC CCCTGACGAG CATCACAAAA 650
ATCGACGCTC AAGTCAGAGG TGGCGAAACC CGACAGGACT ATAAAGATAC 700
CAGGCGTTTC CCCCTGGAAG CTCCCTCGTG CGCTCTCCTG TTCCGACCCT 750
GCCGCTTACC GGATACCTGT CCGCCTTTCT CCCTTCGGGA AGCGTGGCGC 800
TTTCTCATAG CTCACGCTGT AGGTATCTCA GTTCGGTGTA GGTCGTTCGC 850
TCCAAGCTGG GCTGTGTGCA CGAACCCCCC GTTCAGCCCG ACCGCTGCGC 900
CTTATCCGGT AACTATCGTC TTGAGTCCAA CCCGGTAAGA CACGACTTAT 950
CGCCACTGGC AGCAGCCACT GGTAACAGGA TTAGCAGAGC GAGGTATGTA 1000
GGCGGTGCTA CAGAGTTCTT GAAGTGGTGG CCTAACTACG GCTACACTAG 1050
AAGGACAGTA TTTGGTATCT GCGCTCTGCT GAAGCCAGTT ACCTTCGGAA 1100
AAAGAGTTGG TAGCTCTTGA TCCGGCAAAC AAACCACCGC TGGTAGCGGT 1150
GGTTTTTTTG TTTGCAAGCA GCAGATTACG CGCAGAAAAA AAGGATCTCA 1200
AGAAGATCCT TTGATCTTTT CTACGGGGTC TGACGCTCAG TGGAACGAAA 1250
ACTCACGTTA AGGGATTTTG GTCATGAGAT TATCAAAAAG GATCTTCACC 1300
TAGATCCTTT TAAATTAAAA ATGAAGTTTT AAATCAATCT AAAGTATATA 1350
TGAGTAAACT TGGTCTGACA GTTACCAATG CTTAATCAGT GAGGCACCTA 1400
TCTCAGCGAT CTGTCTATTT CGTTCATCCA TAGTTGCCTG ACTCCCCGTC 1450
GTGTAGATAA CTACGATACG GGAGGGCTTA CCATCTGGCC CCAGTGCTGC 1500
AATGATACCG CGAGACCCAC GCTCACCGGC TCCAGATTTA TCAGCAATAA 1550
ACCAGCCAGC CGGAAGGGCC GAGCGCAGAA GTGGTCCTGC AACTTTATCC 1600
GCCTCCATCC AGTCTATTAA TTGTTGCCGG GAAGCTAGAG TAAGTAGTTC 1650
GCCAGTTAAT AGTTTGCGCA ACGTTGTTGC CATTGCTACA GGCATCGTGG 1700
TGTCACGCTC GTCGTTTGGT ATGGCTTCAT TCAGCTCCGG TTCCCAACGA 1750
TCAAGGCGAG TTACATGATC CCCCATGTTG TGCAAAAAAG CGGTTAGCTC 1800
CTTCGGTCCT CCGATCGTTG TCAGAAGTAA GTTGGCCGCA GTGTTATCAC 1850
TCATGGTTAT GGCAGCACTG CATAATTCTC TTACTGTCAT GCCATCCGTA 1900
AGATGCTTTT CTGTGACTGG TGAGTACTCA ACCAAGTCAT TCTGAGAATA 1950

GTGTATGCGG CGACCGAGTT GCTCTTGCCC GGCGTCAACA CGGGATAATA 2000
CCGCGCCACA TAGCAGAACT TTAAAAGTGC TCATCATTGG AAAACGTTCT 2050
TCGGGGCGAA AACTCTCAAG GATCTTACCG CTGTTGAGAT CCAGTTCGAT 2100
GTAACCCACT CGTGCACCCA ACTGATCTTC AGCATCTTTT ACTTTCACCA 2150
GCGTTTCTGG GTGAGCAAAA ACAGGAAGGC AAAATGCCGC AAAAAAGGGA 2200
ATAAGGGCGA CACGGAAATG TTGAATACTC ATACTCTTCC TTTTTCAATA 2250
TTATTGAAGC ATTTATCAGG GTTATTGTCT CATGAGCGGA TACATATTTG 2300
AATGTATTTA GAAAAATAAA CAAATAGGGG TTCCGCGCAC ATTTCCCCGA 2350
AAAGTGCCAC CTGACGTCTA AGAAACCATT ATTATCATGA CATTAACCTA 2400
TAAAAATAGG CGTATCACGA GGCCCTTTCG TCTTCAAGAA TTCTGAACCA 2450
GTCCTAAAAC GAGTAAATAG GACCGGCAAT TCTTCAAGCA ATAAACAGGA 2500
ATACCAATTA TTAAAAGATA ACTTAGTCAG ATCGTACAAT AAAGCTAGCT 2550
TTGAAGAAAA ATGCGCCTTA TTCAATCTTT GCTATAAAAA ATGGCCCAAA 2600
ATCTCACATT GGAAGACATT TGATGACCTC ATTTCTTTCA ATGAAGGGCC 2650
TAACGGAGTT GACTAATGTT GTGGGAAATT GGAGCGATAA GCGTGCTTCT 2700
GCCGTGGCCA GGACAACGTA TACTCATCAG ATAACAGCAA TACCTGATCA 2750
CTACTTCGCA CTAGTTTCTC GGTACTATGC ATATGATCCA ATATCAAAGG 2800
AAATGATAGC ATTGAAGGAT GAGACTAATC CAATTGAGGA GTGGCAGCAT 2850
ATAGAACAGC TAAAGGGTAG TGCTGAAGGA AGCATACGAT ACCCCGCATG 2900
GAATGGGATA ATATCACAGG AGGTACTAGA CTACCTTTCA TCCTACATAA 2950
ATAGACGCAT ATAAGTACGC ATTTAAGCAT AAACACGCAC TATGCCGTTC 3000
TTCTCATGTA TATATATATA CAGGCAACAC GCAGATATAG GTGCGACGTG 3050
AACAGTGAGC TGTATGTGCG CAGCTCGCGT TGCATTTTCG GAAGCGCTCG 3100
TTTTCGGAAA CGCTTTGAAG TTCCTATTCC GAAGTTCCTA TTCTCTAGAA 3150
AGTATAGGAA CTTCAGAGCG CTTTTGAAAA CCAAAAGCGC TCTGAAGACG 3200
CACTTTCAAA AAACCAAAAA CGCACCGGAC TGTAACGAGC TACTAAAATA 3250
TTGCGAATAC CGCTTCCACA AACATTGCTC AAAAGTATCT CTTTGCTATA 3300
TATCTCTGTG CTATATCCCT ATATAACCTA CCCATCCACC TTTCGCTCCT 3350
TGAACTTGCA TCTAAACTCG ACCTCTACAT TTTTTATGTT TATCTCTAGT 3400
ATTACTCTTT AGACAAAAAA ATTGTAGTAA GAACTATTCA TAGAGTGAAT 3450
CGAAAACAAT ACGAAAATGT AAACATTTCC TATACGTAGT ATATAGAGAC 3500
AAAATAGAAG AAACCGTTCA TAATTTTCTG ACCAATGAAG AATCATCAAC 3550
GCTATCACTT TCTGTTCACA AAGTATGCGC AATCCACATC GGTATAGAAT 3600
ATAATCGGGG ATGCCTTTAT CTTGAAAAAA TGCACCCGCA GCTTCGCTAG 3650
TAATCAGTAA ACGCGGGAAG TGGAGTCAGG CTTTTTTTAT GGAAGAGAAA 3700
ATAGACACCA AAGTAGCCTT CTTCTAACCT TAACGGACCT ACAGTGCAAA 3750
AAGTTATCAA GAGACTGCAT TATAGAGCGC ACAAAGGAGA AAAAAAGTAA 3800
TCTAAGATGC TTTGTTAGAA AAATAGCGCT CTCGGGATGC ATTTTTGTAG 3850
AACAAAAAAG AAGTATAGAT TCTTTGTTGG TAAAATAGCG CTCTCGCGTT 3900
GCATTTCTGT TCTGTAAAAA TGCAGCTCAG ATTCTTTGTT TGAAAAATTA 3950
GCGCTCTCGC GTTGCATTTT TGTTTTACAA AAATGAAGCA CAGATTCTTC 4000
GTTGGTAAAA TAGCGCTTTC GCGTTGCATT TCTGTTCTGT AAAAATGCAG 4050
CTCAGATTCT TTGTTTGAAA AATTAGCGCT CTCGCGTTGC ATTTTTGTTC 4100
TACAAAATGA AGCACAGATG CTTCGTTAAC AAAGATATGC TATTGAAGTG 4150
CAAGATGGAA ACGCAGAAAA TGAACCGGGG ATGCGACGTG CAAGATTACC 4200
TATGCAATAG ATGCAATAGT TTCTCCAGGA ACCGAAATAC ATACATTGTC 4250
TTCCGTAAAG CGCTAGACTA TATATTATTA TACAGGTTCA AATATACTAT 4300
CTGTTTCAGG GAAAACTCCC AGGTTCGGAT GTTCAAAATT CAATGATGGG 4350
TAACAAGTAC GATCGTAAAT CTGTAAAACA GTTTGTCGGA TATTAGGCTG 4400

TATCTCCTCA AAGCGTATTC GAATATCATT GAGAAGCTGC ATTTTTTTTT 4450
TTTTTTATAT ATATTTCAAG GATATACCAT TGTAATGCCT GCCCCTAAGA 4500
AGATCGTCGT TTTGCCAGGT GACCACGTTG GTCAAGAAAT CACAGCCGAA 4550
GCCATTAAGG TTCTTAAAGC TATTTCTGAT GTTCGTTCCA ATGTCAAGTT 4600
CGATTTCGAA AATCATTTAA TTGGTGGTGC TGCTATCGAT GCTACAGGTG 4650
TTCCACTTCC AGATGAGGCG CTGGAAGCCT CCAAGAAGGC TGATGCCGTT 4700
TTGTTAGGTG CTGTGGGTGG TCCTAAATGG GGTACCGGTA GTGTTAGACC 4750
TGAACAAGGT TTACTAAAAA TCCGTAAAGA ACTTCAATTG TACGCCAACT 4800
TAAGACCATG TAACTTTGCA TCCGACTCTC TTTTAGACTT ATCTCCAATC 4850
AAGCCACAAT TTGCTAAAGG TACTGACTTC GTTGTTGTTA GAGAATTAGT 4900
GGGAGGTATT TACTTTGGTA AGAGAAAGGA AGACGATGGT GATGGTGTCG 4950
CTTGGGATAG TGAACAATAC ACCGTTCCAG AAGTGCAAAG AATCACAAGA 5000
ATGGCCGCTT TCATGGCCCT ACAACATGAG CCACCATTGC CTATTTGGTC 5050
CTTGGATAAA GCTAATGTTT TGGCCTCTTC AAGATTATGG AGAAAAACTG 5100
TGGAGGAAAC CATCAAGAAC GAATTCCCTA CATTGAAAGT TCAACATCAA 5150
TTGATTGATT CTGCCGCCAT GATCCTAGTT AAGAACCCAA CCCACCTAAA 5200
TGGTATTATA ATCACCAGCA ACATGTTTGG TGATATCATC TCCGATGAAG 5250
CCTCCGTTAT CCCAGGCTCC TTGGGTTTGT TGCCATCTGC GTCCTTGGCC 5300
TCTTTGCCAG ACAAGAACAC CGCATTTGGT TTGTACGAAC CATGCCATGG 5350
TTCCGCTCCA GATTTGCCAA AGAATAAGGT CAACCCTATC GCCACTATCT 5400
TGTCTGCTGC AATGATGTTG AAATTGTCAT TGAACTTGCC TGAAGAAGGT 5450
AAAGCCATTG AAGATGCAGT TAAAAAGGTT TTGGATGCAG GTATCAGAAC 5500
TGGTGATTTA GGTGGTTCCA ACAGTACCAC CGAAGTCGGT GATGCTGTCG 5550
CCGAAGAAGT TAAGAAAATC CTTGCTTAAA AAGATTCTCT TTTTTTATGA 5600
TATTTGTACA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 5650
AAAAAAAAAA AAAATGCAGC GTCACATCGG ATAATAATGA TGGCAGCCAT 5700
TGTAGAAGTG CCTTTTGCAT TTCTAGTCTC TTTCTCGGTC TAGCTAGTTT 5750
TACTACATCG CGAAGATAGA ATCTTAGATC ACACTGCCTT TGCTGAGCTG 5800
GATCAATAGA GTAACAAAAG AGTGGTAAGG CCTCGTTAAA GGACAAGGAC 5850
CTGAGCGGAA GTGTATCGTA CAGTAGACGG AGTATACTAG TATAGTCTAT 5900
AGTCCGTGGA ATTCTCATGT TTGACAGCTT ATCATCGATA AGCTAGCTTT 5950
CTAACTGATC TATCCAAAAC TGAAAATTAC ATTCTTGATT AGGTTTATCA 6000
CAGGCAAATG TAATTTGTGG TATTTTGCCG TTCAAAATCT GTAGAATTTT 6050
CTCATTGGTC ACATTACAAC CTGAAAATAC TTTATCTACA ATCATACCAT 6100
TCTTAATAAC ATGTCCCCTT AATACTAGGA TCAGGCATGA ACGCATCACA 6150
GACAAAATCT TCTTGACAAA CGTCACAATT GATCCCTCCC CATCCGTTAT 6200
CACAATGACA GGTGTCATTT TGTGCTCTTA TGGGACGATC CTTATTACCG 6250
CTTTCATCCG GTGATTGACC GCCACAGAGG GGCAGAGAGC AATCATCACC 6300
TGCAAACCCT TCTATACACT CACATCTACC AGTGATCGAA TTGCATTCAG 6350
AAAACTGTTT GCATTCAAAA ATAGGTAGCA TACAATTAAA ACATGGCGGG 6400
CATGTATCAT TGCCCTTATC TTGTGCAGTT AGACGCGAAT TTTTCGAAGA 6450
AGTACCTTCA AAGAATGGGG TCTTATCTTG TTTTGCAAGT ACCACTGAGC 6500
AGGATAATAA TAGAAATGAT AATATACTAT AGTAGAGATA ACGTCGATGA 6550
CTTCCCATAC TGTAATTGCT TTTAGTTGTG TATTTTTAGT GTGCAAGTTT 6600
CTGTAAATCG ATTAATTTTT TTTTCTTTCC TCTTTTTATT AACCTTAATT 6650
TTTATTTTAG ATTCCTGACT TCAACTCAAG ACGCACAGAT ATTATAACAT 6700
CTGCATAATA GGCATTTGCA AGAATTACTC GTGAGTAAGG AAAGAGTGAG 6750
GAACTATCGC ATACCTGCAT TTAAAGATGC CGATTTGGGC GCGAATCCTT 6800
TATTTTGGCT TCACCCTCAT ACTATTATCA GGGCCAGAAA AAGGAAGTGT 6850

```
TTCCCTCCTT CTTGAATTGA TGTTACCCTC ATAAAGCACG TGGCCTCTTA 6900
TCGAGAAAGA AATTACCGTC GCTCGTGATT TGTTTGCAAA AAGAACAAAA 6950
CTGAAAAAAC CCAGACACGC TCGACTTCCT GTCTTCCTAT TGATTGCAGC 7000
TTCCAATTTC GTCACACAAC AAGGTCCTAG CGACGGCTCA CAGGTTTTGT 7050
AACAAGCAAT CGAAGGTTCT GGAATGGCGG GGAAAGGGTT TAGTACCACA 7100
TGCTATGATG CCCACTGTGA TCTCCAGAGC AAAGTTCGTT CGATCGTACT 7150
GTACTCTCTC TCTTTCAAAC AGAATTGTCC GAATCGTGTG ACAACAACAG 7200
CCTGTTCTCA CACACTCTTT TCTTCTAACC AAGGGGGTGG TTTAGTTTAG 7250
TAGAACCTCG TGAAACTTAC ATTTACATAT ATATAAACTT GCATAAATTG 7300
GTCAATGGAA GAAATACATA TTTGGTCTTT TCTAATTCGT AGTTTTTCAA 7350
GTTCTTAGAT GCTTTCTTTT TCTCTTTTTT ACAGATCATC AAGGAAGTAA 7400
TTATCTACTT TTTACAACAA ATACAAAAGA TCTATGAGAT TTCCTTCAAT 7450
TTTTACTGCA GTTTTATTCG CAGCATCCTC CGCATTAGCT GCTCCAGTCA 7500
ACACTACAAC AGAAGATGAA ACGGCACAAA TTCCGGCTGA AGCTGTCATC 7550
GGTTACTTAG ATTTAGAAGG GGATTTCGAT GTTGCTGTTT TGCCATTTTC 7600
CAACAGCACA AATAACGGGT TATTGTTTAT AAATACTACT ATTGCCAGCA 7650
TTGCTGCTAA AGAAGAAGGG GTAAGCTTGG ATAAAAGAAA CAGCGACTCT 7700
GAATGCCCGC TGAGCCATGA TGGCTACTGC CTGCACGACG GTGTATGCAT 7750
GTATATCGAA GCTCTGGACA AATACGCATG CAACTGCGTA GTTGGTTACA 7800
TCGGCGAACG TTGCCAGTAC CGCGACCTGA AATGGTGGGA GCTCCGTTAA 7850
TAAGGATCC                                              7859
```

**** END OF SEQ ID NO: 4 *****

SEQ. ID NO:5

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 15 base pairs
FEATURES: Top strand of adapter to fuse C-terminal end of the α-factor pro-peptide to synthetic hirudin gene

SEQUENCE:

AGCTTGGATA AAAGA 15

**** END OF SEQ ID NO: 5 *****

SEQ. ID NO:6

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 11 base pairs
FEATURES: Bottom strand of adapter to fuse C-terminal end of the α-factor pro-peptide to synthetic hirudin gene

SEQUENCE:

TCTTTTATCC A 11

**** END OF SEQ ID NO: 6 *****

SEQ ID NO:7

SEQUENCE TYPE:	nucleotide
SEQUENCE LENGTH:	223 base pairs
STRANDEDNESS:	single
TOPOLOGY:	linear
MOLECULE TYPE:	synthetic DNA
SOURCE:	synthetic
FEATURES:	hirudin type HV-1 gene with 5 amino acid adaptor (corresponding to C-terminus of alpha factor) at amino terminus. from 1 to 6 bp (AAGCTT) is HindIII site from 118 to 123 bp (GGATCC) is BamHI site.

SEQUENCE:

```
AAGCTTGGAT AAAAGAGTTG TTTACACCGA CTGTACTGAA TCCGGACAAA   50
ACCTGTGTTT GTGTGAGGGT TCTAACGTCT GTGGTCAGGG TAACAAATGC  100
ATCCTGGGTT CCGACGGTGA AAAGAACCAA TGTGTCACTG GTGAAGGTAC  150
CCCAAAGCCG CAGTCCCACA ACGATGGAGA TTTCGAAGAA ATCCCAGAAG  200
AATATCTGCA GTAATAGGGA TCC                               223
```

**** END OF SEQ ID NO: 7 *****

SEQ ID NO:8
SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 420 base pairs
STRANDEDNESS: double
TOPOLOGY: linear
FEATURES: Factor Xa-cleavable Hirudin-IEGR-Hirudin
SEQUENCE:

```
GTT GTT TAC ACC GAC TGT ACT GAA TCC GGA CAA AAC CTG TGT      42
Val Val Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Leu Cys
                  5                  10

TTG TGT GAG GGT TCT AAC GTC TGT GGT CAG GGT AAC AAA TGC      84
Leu Cys Glu Gly Ser Asn Val Cys Gly Gln Gly Asn Lys Cys
 15                  20                  25

ATC CTG GGT TCC GAC GGT GAA AAG AAC CAA TGT GTC ACT GGT     126
Ile Leu Gly Ser Asp Gly Glu Lys Asn Gln Cys Val Thr Gly
      30                  35                  40

GAA GGT ACC CCA AAG CCG CAG TCC CAC AAC GAT GGA GAT TTC     168
Glu Gly Thr Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe
          45                  50                  55

GAA GAA ATC CCA GAA GAA TAT CTG CAG ATC GAA GGA AGA GTT     210
Glu Glu Ile Pro Glu Glu Tyr Leu Gln Ile Glu Gly Arg Val
              60                  65                  70

GTT TAC ACC GAC TGT ACT GAA TCC GGA CAA AAC CTG TGT TTG     252
Val Tyr Thr Asp Cys Thr Glu Ser Gly Gln Asn Leu Cys Leu
                  75                  80

TGT GAG GGT TCT AAC GTC TGT GGT CAG GGT AAC AAA TGC ATC     294
Cys Glu Gly Ser Asn Val Cys Gly Gln Gly Asn Lys Cys Ile
 85                  90                  95

CTG GGT TCC GAC GGT GAA AAG AAC CAA TGT GTC ACT GGT GAA     336
Leu Gly Ser Asp Gly Glu Lys Asn Gln Cys Val Thr Gly Glu
     100                 105                 110

GGT ACC CCA AAG CCG CAG TCC CAC AAC GAT GGA GAT TTC GAA     378
Gly Thr Pro Lys Pro Gln Ser His Asn Asp Gly Asp Phe Glu
         115                 120                 125

GAA ATC CCA GAA GAA TAT CTG CAG TAATAGGGAT CCGAATTC         420
Glu Ile Pro Glu Glu Tyr Leu Gln
             130
```

**** END OF SEQ ID NO: 8 *****

SEQ. ID NO:13

SEQUENCE TYPE: nucleotide
SEQUENCE LENGTH: 17 base pairs
FEATURES: Primers for dideoxy sequencing of streptokinase gene

SEQUENCE:

5'-CACTATCAGTAGCAAAT-3'	BB 3510
5'-TGGTCTAACGCGCACAT-3'	BB 2136
5'-GAGTAAACTGTACAGTA-3'	BB 3509
5'-GATCTCATAAGCTTGTT-3'	BB 3508
5'-TTTAGCCTTATCACGAG-3'	BB 2135
5'-GACACCAACCGTATCAT-3'	BB 2753
5'-CGTTGATGTCAACACCA-3'	BB 3718
5'-GCTATCGGTGACACCAT-3'	BB 2754
5'-GACGACTACTTTGAGGT-3'	BB 2755
5'-CCCAACCTGTCCAAGAA-3'	BB 2134

**** END OF SEQ ID NO: 13 *****

SEQ. ID NO:14

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1335 base pairs
FEATURES: Streptokinase gene from S. equisimilis
SEQUENCE:

```
GAATTCATGAAAAATTACTTATCTTTTGGGATGTTTGCACTGCTGTTTGCACTAACATTT
      MetLysAsnTyrLeuSerPheGlyMetPheAlaLeuLeuPheAlaLeuThrPhe

GGAACAGTCAATTCTGTCCAAGCTATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCT
GlyThrValAsnSerValGlnAlaIleAlaGlyProGluTrpLeuLeuAspArgProSer

GTCAACAACAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGAC
ValAsnAsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAsp

ATTAGTCTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACA
IleSerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThr

GAGCAAGGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACAT
GluGlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHis

AAACTTGAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCAC
LysLeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHis

AGTAACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGA
SerAsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArg

AACGGCAAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTC
AsnGlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProVal

CAAGAATTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAA
GlnGluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGln

AATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGAT
AsnGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAsp

GACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGAC
AspAspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAsp

ACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCA
ThrIleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisPro

GGCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGT
GlyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArg

ACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTAT
ThrIleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyr
```

GAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAG
GluIleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGlu

AAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTG
LysTyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeu

AAACTGTTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAG
LysLeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGln

CTCTTAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAG
LeuLeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLys

GCTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGA
AlaLysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGly

AAAGTAGAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGA
LysValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArg

CCCGAAGGAGAGAATGCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAA
ProGluGlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGlu

GAACGAGAAGTTTACAGCTACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAAC
GluArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsn

GACAAATAAGGATCC*
AspLysEnd

**** END OF SEQ ID NO: 14 *****

SEQ. ID NO:17

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1317 base pairs
FEATURES: OmpAL fused to mature streptokinase gene
SEQUENCE:

```
CATATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCGACCGTAGCG
  M  K  K  T  A  I  A  I  A  V  A  L  A  G  F  A  T  V  A

CAGGCCATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTA
Q  A  I  A  G  P  E  W  L  L  D  R  P  S  V  N  N  S  Q  L

GTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTT
V  V  S  V  A  G  T  V  E  G  T  N  Q  D  I  S  L  K  F  F

GAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCA
E  I  D  L  T  S  R  P  A  H  G  G  K  T  E  Q  G  L  S  P

AAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGAC
K  S  K  P  F  A  T  D  S  G  A  M  P  H  K  L  E  K  A  D

TTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTT
L  L  K  A  I  Q  E  Q  L  I  A  N  V  H  S  N  D  D  Y  F

GAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTT
E  V  I  D  F  A  S  D  A  T  I  T  D  R  N  G  K  V  Y  F

GCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGC
A  D  K  D  G  S  V  T  L  P  T  Q  P  V  Q  E  F  L  L  S

GGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTT
G  H  V  R  V  R  P  Y  K  E  K  P  I  Q  N  Q  A  K  S  V

GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGT
D  V  E  Y  T  V  Q  F  T  P  L  N  P  D  D  D  F  R  P  G

CTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAA
L  K  D  T  K  L  L  K  T  L  A  I  G  D  T  I  T  S  Q  E

TTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATGAA
L  L  A  Q  A  Q  S  I  L  N  K  T  H  P  G  Y  T  I  Y  E

CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGAT
R  D  S  S  I  V  T  H  D  N  D  I  F  R  T  I  L  P  M  D

CAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCT
Q  E  F  T  Y  H  V  K  N  R  E  Q  A  Y  E  I  N  K  K  S
```

```
GGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAA
G  L  N  E  E  I  N  N  T  D  L  I  S  E  K  Y  Y  V  L  K

AAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAA
K  G  E  K  P  Y  D  P  F  D  R  S  H  L  K  L  F  T  I  K

TACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAA
Y  V  D  V  N  T  N  E  L  L  K  S  E  Q  L  L  T  A  S  E

CGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAC
R  N  L  D  F  R  D  L  Y  D  P  R  D  K  A  K  L  L  Y  N

AATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAAGATAATCAC
N  L  D  A  F  G  I  M  D  Y  T  L  T  G  K  V  E  D  N  H

GATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCT
D  D  T  N  R  I  I  T  V  Y  M  G  K  R  P  E  G  E  N  A

AGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGC
S  Y  H  L  A  Y  D  K  D  R  Y  T  E  E  E  R  E  V  Y  S

TACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAAGGATCC*
Y  L  R  Y  T  G  T  P  I  P  D  N  P  N  D  K  *
```

**** END OF SEQ ID NO: 17 *****

SEQ. ID NO:23

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1197 nucleotides
FEATURES: Methionyl-streptokinase fusion protein
SEQUENCE:

```
CATATGATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTA
   MetIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeu

GTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTT
ValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePhe

GAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCA
GluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerPro

AAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGAC
LysSerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAsp

TTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTT
LeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPhe

GAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTT
GluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPhe

GCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGC
AlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSer

GGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTT
GlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerVal

GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGT
AspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGly

CTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAA
LeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlu

TTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATGAA
LeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlu

CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGAT
ArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAsp

CAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCT
GlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSer

GGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAA
GlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLys
```

```
AAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAA
LysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLys

TACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAA
TyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGlu

CGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAC
ArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsn

AATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAAGATAATCAC
AsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHis

GATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCT
AspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAla

AGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGC
SerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSer

TACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAATAAGGATCC*
TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysEnd
```

**** END OF SEQ ID NO: 23 *****

SEQ. ID NO:24

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1513 nucleotides
FEATURES: Streptokinase fused to yeast α-factor
SEQUENCE:

```
AGATCTATGAGATTTCCTTCAATTTTTACTGCAGTTTTATTCGCAGCATCCTCCGCATTA
      MetArgPheProSerIlePheThrAlaValLeuPheAlaAlaSerSerAlaLeu

GCTGCTCCAGTCAACACTACAACAGAAGATGAAACGGCACAAATTCCGGCTGAAGCTGTC
AlaAlaProValAsnThrThrThrGluAspGluThrAlaGlnIleProAlaGluAlaVal

ATCGGTTACTTAGATTTAGAAGGGGATTTCGATGTTGCTGTTTTGCCATTTTCCAACAGC
IleGlyTyrLeuAspLeuGluGlyAspPheAspValAlaValLeuProPheSerAsnSer

ACAAATAACGGGTTATTGTTTATAAATACTACTATTGCCAGCATTGCTGCTAAAGAAGAA
ThrAsnAsnGlyLeuLeuPheIleAsnThrThrIleAlaSerIleAlaAlaLysGluGlu

GGGGTAAGCTTGGATAAAAGAATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTC
GlyValSerLeuAspLysArgIleAlaGlyProGluTrpLeuLeuAspArgProSerVal

AACAACAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATT
AsnAsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle

AGTCTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAG
SerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGlu

CAAGGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAA
GlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLys

CTTGAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGT
LeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSer

AACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAAC
AsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn

GGCAAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAA
GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln

GAATTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAT
GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsn

CAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGAC
GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAsp

GATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACC
AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr
```

ATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGC
IleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly

TATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACG
TyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr

ATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAG
IleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGlu

ATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAA
IleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLys

TATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAA
TyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLys

CTGTTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTC
LeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeu

TTAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCT
LeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla

AAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAA
LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys

GTAGAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCC
ValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro

GAAGGAGAGAATGCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAA
GluGlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGlu

CGAGAAGTTTACAGCTACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGAC
ArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAsp

AAATAAGGATCC*
LysEnd

**** END OF SEQ ID NO: 24 *****

SEQ. ID NO:26

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1120 nucleotides
FEATURES: Truncated Met-streptokinase (aa 16-383)
SEQUENCE:

```
CATATGAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATT
   MetSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle

AGTCTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAG
SerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGlu

CAAGGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAA
GlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLys

CTTGAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGT
LeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSer

AACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAAC
AsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn

GGCAAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAA
GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln

GAATTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAT
GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsn

CAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGAC
GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAsp

GATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACC
AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr

ATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGC
IleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly

TATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACG
TyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr

ATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAG
IleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGlu

ATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAA
IleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLys

TATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAA
TyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLys
```

```
CTGTTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTC
LeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeu

TTAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCT
LeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla

AAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAA
LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys

GTAGAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCC
ValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro

GAAGGAGAGAATGCTAGCTATCATTTAGCCTAAGGATCC*
GluGlyGluAsnAlaSerTyrHisLeuAlaEnd
```

**** END OF SEQ ID NO: 26 *****

SEQ. ID NO:29

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 2590 nucleotides
FEATURES: OmpAL-Streptokinase-streptokinase fusion linked by thrombin-cleavable VELQGVVPRG

SEQUENCE:

```
CATATGAAAAAGACAGCTATCGCGATTGCAGTGGCACTGGCTGGTTTCGCGACCGTAGCG
   MetLysLysThrAlaIleAlaIleAlaValAlaLeuAlaGlyPheAlaThrValAla

CAGGCCATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTA
GlnAlaIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeu

GTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTT
ValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePhe

GAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCA
GluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerPro

AAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGAC
LysSerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAsp

TTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTT
LeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPhe

GAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTT
GluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPhe

GCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGC
AlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSer

GGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTT
GlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerVal

GATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGT
AspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGly

CTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAA
LeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGlu

TTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATGAA
LeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGlu

CGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGAT
ArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAsp

CAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCT
GlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSer
```

GGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAA
GlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLys

AAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAA
LysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLys

TACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAA
TyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGlu

CGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAAC
ArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsn

AATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAAGATAATCAC
AsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHis

GATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCT
AspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAla

AGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGC
SerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSer

TACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAAGTAGAGCTGCAG
TyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysValGluLeuGln

GGAGTAGTTCCTCGTGGAATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAAC
GlyValValProArgGlyIleAlaGlyProGluTrpLeuLeuAspArgProSerValAsn

AACAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGT
AsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSer

CTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAA
LeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGln

GGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTT
GlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeu

GAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAAC
GluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsn

GACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGC
AspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGly

AAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAA
LysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGlu

TTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAA
PheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGln

GCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGAT
AlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAsp

TTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATC
PheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIle

ACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTAT
ThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyr

ACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATT
ThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIle

TTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATC
LeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIle

AATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATAT
AsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyr

TACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTG
TyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeu

TTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTA
PheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeu

ACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAA
ThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLys

CTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTA
LeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysVal

GAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAA
GluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGlu

GGAGAGAATGCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGA
GlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArg

GAAGTTTACAGCTACCTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAA
GluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLys

TAAGGATCC*
End

**** END OF SEQ ID NO: 29 *****

SEQ. ID NO:33

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 2254 nucleotides
FEATURES: Met-corestreptokinase-corestreptokinase fusion linked by thrombin-cleavable VELQGVVPRG

SEQUENCE:

```
CATATGAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATT
   MetSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIle

AGTCTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAG
SerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGlu

CAAGGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAA
GlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLys

CTTGAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGT
LeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSer

AACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAAC
AsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsn

GGCAAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAA
GlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGln

GAATTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAAT
GluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsn

CAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGAC
GlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAsp

GATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACC
AspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThr

ATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGC
IleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGly

TATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACG
TyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThr

ATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAG
IleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGlu

ATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAA
IleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLys
```

```
TATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAA
TyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLys

CTGTTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTC
LeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeu

TTAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCT
LeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAla

AAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAA
LysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLys

GTAGAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCC
ValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgPro

GAAGGAGAGAATGCTAGCTATCATTTAGCCGTAGAGCTGCAGGGAGTAGTTCCTCGTGGA
GluGlyGluAsnAlaSerTyrHisLeuAlaValGluLeuGlnGlyValValProArgGly

AGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTT
SerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeu

AAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGC
LysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGly

TTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAA
LeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGlu

AAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGAC
LysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAsp

GACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAG
AspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLys

GTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTT
ValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnProValGlnGluPhe

TTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCG
LeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAla

AAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTC
LysSerValAspValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPhe

AGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACA
ArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThr

TCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACG
SerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThr
```

ATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTA
IleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeu

CCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAAT
ProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsn

AAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTAC
LysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyr

GTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTC
ValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPhe

ACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACA
ThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThr

GCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTA
AlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeu

CTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAA
LeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGlu

GATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGA
AspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGly

GAGAATGCTAGCTATCATTTAGCCTAAGGATCC
GluAsnAlaSerTyrHisLeuAlaEnd

**** END OF SEQ ID NO: 33 *****

SEQ. ID NO:35

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1459 nucleotides
FEATURES: Hirudin-streptokinase fusion linked by Factor Xa-cleavable IEGR
SEQUENCE:

```
GTTGTTTACACCGACTGTACTGAATCCGGACAAAACCTGTGTTTGTGTGAGGGTTCTAAC
ValValTyrThrAspCysThrGluSerGlyGlnAsnLeuCysLeuCysGluGlySerAsn

GTCTGTGGTCAGGGTAACAAATGCATCCTGGGTTCCGACGGTGAAAAGAACCAATGTGTC
ValCysGlyGlnGlyAsnLysCysIleLeuGlySerAspGlyGluLysAsnGlnCysVal

ACTGGTGAAGGTACCCCAAAGCCGCAGTCCCACAACGATGGAGATTTCGAAGAAATCCCA
ThrGlyGluGlyThrProLysProGlnSerHisAsnAspGlyAspPheGluGluIlePro

GAAGAATATCTGCAGATCGAAGGTAGAATTGCTGGACCTGAGTGGCTGCTAGACCGTCCA
GluGluTyrLeuGlnIleGluGlyArgIleAlaGlyProGluTrpLeuLeuAspArgPro

TCTGTCAACAACAGCCAATTAGTTGTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAA
SerValAsnAsnSerGlnLeuValValSerValAlaGlyThrValGluGlyThrAsnGln

GACATTAGTCTTAAATTTTTTGAAATTGACCTAACATCACGACCTGCTCATGGAGGAAAG
AspIleSerLeuLysPhePheGluIleAspLeuThrSerArgProAlaHisGlyGlyLys

ACAGAGCAAGGCTTAAGTCCAAAATCAAAACCATTTGCTACTGATAGTGGCGCGATGCCA
ThrGluGlnGlyLeuSerProLysSerLysProPheAlaThrAspSerGlyAlaMetPro

CATAAACTTGAAAAAGCTGACTTACTAAAGGCTATTCAAGAACAATTGATCGCTAACGTC
HisLysLeuGluLysAlaAspLeuLeuLysAlaIleGlnGluGlnLeuIleAlaAsnVal

CACAGTAACGACGACTACTTTGAGGTCATTGATTTTGCAAGCGATGCAACCATTACTGAT
HisSerAsnAspAspTyrPheGluValIleAspPheAlaSerAspAlaThrIleThrAsp

CGAAACGGCAAGGTCTACTTTGCTGACAAAGATGGTTCGGTAACCTTGCCGACCCAACCT
ArgAsnGlyLysValTyrPheAlaAspLysAspGlySerValThrLeuProThrGlnPro

GTCCAAGAATTTTTGCTAAGCGGACATGTGCGCGTTAGACCATATAAAGAAAAACCAATA
ValGlnGluPheLeuLeuSerGlyHisValArgValArgProTyrLysGluLysProIle

CAAAATCAAGCGAAATCTGTTGATGTGGAATATACTGTACAGTTTACTCCCTTAAACCCT
GlnAsnGlnAlaLysSerValAspValGluTyrThrValGlnPheThrProLeuAsnPro

GATGACGATTTCAGACCAGGTCTCAAAGATACTAAGCTATTGAAAACACTAGCTATCGGT
AspAspAspPheArgProGlyLeuLysAspThrLysLeuLeuLysThrLeuAlaIleGly
```

GACACCATCACATCTCAAGAATTACTAGCTCAAGCACAAAGCATTTTAAACAAAACCCAT
AspThrIleThrSerGlnGluLeuLeuAlaGlnAlaGlnSerIleLeuAsnLysThrHis

CCAGGCTATACGATTTATGAACGTGACTCCTCAATCGTCACTCATGACAATGACATTTTC
ProGlyTyrThrIleTyrGluArgAspSerSerIleValThrHisAspAsnAspIlePhe

CGTACGATTTTACCAATGGATCAAGAGTTTACTTACCATGTCAAAAATCGGGAACAAGCT
ArgThrIleLeuProMetAspGlnGluPheThrTyrHisValLysAsnArgGluGlnAla

TATGAGATCAATAAAAAATCTGGTCTGAATGAAGAAATAAACAACACTGACCTGATCTCT
TyrGluIleAsnLysLysSerGlyLeuAsnGluGluIleAsnAsnThrAspLeuIleSer

GAGAAATATTACGTCCTTAAAAAAGGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCAC
GluLysTyrTyrValLeuLysLysGlyGluLysProTyrAspProPheAspArgSerHis

TTGAAACTGTTCACCATCAAATACGTTGATGTCAACACCAACGAATTGCTAAAAAGCGAG
LeuLysLeuPheThrIleLysTyrValAspValAsnThrAsnGluLeuLeuLysSerGlu

CAGCTCTTAACAGCTAGCGAACGTAACTTAGACTTCAGAGATTTATACGATCCTCGTGAT
GlnLeuLeuThrAlaSerGluArgAsnLeuAspPheArgAspLeuTyrAspProArgAsp

AAGGCTAAACTACTCTACAACAATCTCGATGCTTTTGGTATTATGGACTATACCTTAACT
LysAlaLysLeuLeuTyrAsnAsnLeuAspAlaPheGlyIleMetAspTyrThrLeuThr

GGAAAAGTAGAAGATAATCACGATGACACCAACCGTATCATAACCGTTTATATGGGCAAG
GlyLysValGluAspAsnHisAspAspThrAsnArgIleIleThrValTyrMetGlyLys

CGACCCGAAGGAGAGAATGCTAGCTATCATTTAGCCTATGATAAAGATCGTTATACCGAA
ArgProGluGlyGluAsnAlaSerTyrHisLeuAlaTyrAspLysAspArgTyrThrGlu

GAAGAACGAGAAGTTTACAGCTACCTGCGTTATACAGGGACACCTATACCTGATAACCCT
GluGluArgGluValTyrSerTyrLeuArgTyrThrGlyThrProIleProAspAsnPro

AACGACAAATAAGGATCC*
AsnAspLysEnd

**** END OF SEQ ID NO: 35 *****

SEQ. ID NO:38

SEQUENCE TYPE: nucleotide with corresponding amino acid
SEQUENCE LENGTH: 1468 nucleotides
FEATURES: Streptokinase-hirudin fusion linked by Factor Xa-cleavable IEGR
SEQUENCE:

```
ATTGCTGGACCTGAGTGGCTGCTAGACCGTCCATCTGTCAACAACAGCCAATTAGTT
IleAlaGlyProGluTrpLeuLeuAspArgProSerValAsnAsnSerGlnLeuVal

GTTAGCGTTGCTGGTACTGTTGAGGGGACGAATCAAGACATTAGTCTTAAATTTTTTGAA
ValSerValAlaGlyThrValGluGlyThrAsnGlnAspIleSerLeuLysPhePheGlu

ATTGACCTAACATCACGACCTGCTCATGGAGGAAAGACAGAGCAAGGCTTAAGTCCAAAA
IleAspLeuThrSerArgProAlaHisGlyGlyLysThrGluGlnGlyLeuSerProLys

TCAAAACCATTTGCTACTGATAGTGGCGCGATGCCACATAAACTTGAAAAAGCTGACTTA
SerLysProPheAlaThrAspSerGlyAlaMetProHisLysLeuGluLysAlaAspLeu

CTAAAGGCTATTCAAGAACAATTGATCGCTAACGTCCACAGTAACGACGACTACTTTGAG
LeuLysAlaIleGlnGluGlnLeuIleAlaAsnValHisSerAsnAspAspTyrPheGlu

GTCATTGATTTTGCAAGCGATGCAACCATTACTGATCGAAACGGCAAGGTCTACTTTGCT
ValIleAspPheAlaSerAspAlaThrIleThrAspArgAsnGlyLysValTyrPheAla

GACAAAGATGGTTCGGTAACCTTGCCGACCCAACCTGTCCAAGAATTTTTGCTAAGCGGA
AspLysAspGlySerValThrLeuProThrGlnProValGlnGluPheLeuLeuSerGly

CATGTGCGCGTTAGACCATATAAAGAAAAACCAATACAAAATCAAGCGAAATCTGTTGAT
HisValArgValArgProTyrLysGluLysProIleGlnAsnGlnAlaLysSerValAsp

GTGGAATATACTGTACAGTTTACTCCCTTAAACCCTGATGACGATTTCAGACCAGGTCTC
ValGluTyrThrValGlnPheThrProLeuAsnProAspAspAspPheArgProGlyLeu

AAAGATACTAAGCTATTGAAAACACTAGCTATCGGTGACACCATCACATCTCAAGAATTA
LysAspThrLysLeuLeuLysThrLeuAlaIleGlyAspThrIleThrSerGlnGluLeu

CTAGCTCAAGCACAAAGCATTTTAAACAAAACCCATCCAGGCTATACGATTTATGAACGT
LeuAlaGlnAlaGlnSerIleLeuAsnLysThrHisProGlyTyrThrIleTyrGluArg

GACTCCTCAATCGTCACTCATGACAATGACATTTTCCGTACGATTTTACCAATGGATCAA
AspSerSerIleValThrHisAspAsnAspIlePheArgThrIleLeuProMetAspGln

GAGTTTACTTACCATGTCAAAAATCGGGAACAAGCTTATGAGATCAATAAAAAATCTGGT
GluPheThrTyrHisValLysAsnArgGluGlnAlaTyrGluIleAsnLysLysSerGly
```

```
CTGAATGAAGAAATAAACAACACTGACCTGATCTCTGAGAAATATTACGTCCTTAAAAAA
LeuAsnGluGluIleAsnAsnThrAspLeuIleSerGluLysTyrTyrValLeuLysLys

GGGGAAAAGCCGTATGATCCCTTTGATCGCAGTCACTTGAAACTGTTCACCATCAAATAC
GlyGluLysProTyrAspProPheAspArgSerHisLeuLysLeuPheThrIleLysTyr

GTTGATGTCAACACCAACGAATTGCTAAAAAGCGAGCAGCTCTTAACAGCTAGCGAACGT
ValAspValAsnThrAsnGluLeuLeuLysSerGluGlnLeuLeuThrAlaSerGluArg

AACTTAGACTTCAGAGATTTATACGATCCTCGTGATAAGGCTAAACTACTCTACAACAAT
AsnLeuAspPheArgAspLeuTyrAspProArgAspLysAlaLysLeuLeuTyrAsnAsn

CTCGATGCTTTTGGTATTATGGACTATACCTTAACTGGAAAAGTAGAAGATAATCACGAT
LeuAspAlaPheGlyIleMetAspTyrThrLeuThrGlyLysValGluAspAsnHisAsp

GACACCAACCGTATCATAACCGTTTATATGGGCAAGCGACCCGAAGGAGAGAATGCTAGC
AspThrAsnArgIleIleThrValTyrMetGlyLysArgProGluGlyGluAsnAlaSer

TATCATTTAGCCTATGATAAAGATCGTTATACCGAAGAAGAACGAGAAGTTTACAGCTAC
TyrHisLeuAlaTyrAspLysAspArgTyrThrGluGluGluArgGluValTyrSerTyr

CTGCGTTATACAGGGACACCTATACCTGATAACCCTAACGACAAAATCGAAGGTAGAGTT
LeuArgTyrThrGlyThrProIleProAspAsnProAsnAspLysIleGluGlyArgVal

GTTTACACCGACTGTACTGAATCCGGACAAAACCTGTGTTTGTGTGAGGGTTCTAACGTC
ValTyrThrAspCysThrGluSerGlyGlnAsnLeuCysLeuCysGluGlySerAsnVal

TGTGGTCAGGGTAACAAATGCATCCTGGGTTCCGACGGTGAAAAGAACCAATGTGTCACT
CysGlyGlnGlyAsnLysCysIleLeuGlySerAspGlyGluLysAsnGlnCysValThr

GGTGAAGGTACCCCAAAGCCGCAGTCCCACAACGATGGAGATTTCGAAGAAATCCCAGAA
GlyGluGlyThrProLysProGlnSerHisAsnAspGlyAspPheGluGluIleProGlu

GAATATCTGCAGTAATAGGGATCCGAATTC*
GluTyrLeuGlnEndEnd
```

**** END OF SEQ ID NO: 38 *****

CLAIMS

1. A fusion protein comprising a first sequence and a second sequence, the fusion protein being cleavable between the first and second sequences by an enzyme involved in blood clotting, wherein after the fusion protein is so cleaved the first and second sequences, or either of them, has greater fibrinolytic and/or anti-thrombotic activity than the uncleaved fusion protein.

2. A fusion protein as claimed in claim 1, which is a cleavable dimer of two fibrinolytic and/or anti-thrombotic proteins.

3. A fusion protein as claimed in claim 1 or 2, wherein the first sequence corresponds to a hirudin or to a protein having the activity of hirudin.

4. A fusion protein as claimed in claim 1 or 2, wherein the first sequence corresponds to streptokinase or to a protein having the activity of streptokinase.

5. A fusion protein as claimed in any one of claims 1 to 4, wherein the second sequence corresponds to a hirudin or to a protein having the activity of hirudin.

6. A fusion protein as claimed in any one of claims 1 to 4, wherein the second sequence corresponds to streptokinase or to a protein having the activity of streptokinase.

7. A fusion protein as claimed in any one of claims 1 to 6, wherein the enzyme involved in blood clotting is kallikrein, Factor XIIa, XIa, IXa, VIIa, Xa, thrombin (Factor IIa) or activated protein C.

8. A fusion protein as claimed in any one of claims 1 to 6, wherein the enzyme involved in blood clotting is Factor Xa or thrombin.

9. A fusion protein as claimed in any one of claims 1 to 6, wherein the enzyme involved in blood clotting is Factor Xa.

10. A fusion protein as claimed in claim 9, comprising the cleavage site sequence P4-P3-Gly-Arg, wherein P4 represents a hydrophobic residue and P3 represents an acidic residue.

11. A fusion protein as claimed in claim 10, wherein the hydrophobic residue is isoleucine.

12. A fusion protein as claimed in any one of claims 1 to 6, wherein the enzyme involved in blood clotting is thrombin.

13. A fusion protein as claimed in claim 12, comprising the cleavage site sequence P4-P3-Pro-Arg-P1'-P2', wherein each of P4 and P3 independently represents a hydrophobic residue and each of P1' and P2' independently represents a non-acidic residue.

14. A fusion protein as claimed in claim 12, comprising the cleavage site sequence P2-Arg-P1', wherein one of the residues P2 and P1' represents glycine, and the other is any amino acid residue.

15. A fusion protein as claimed in claim 12, comprising the cleavage site sequence Gly-Pro-Arg.

16. A process for the preparation of a fusion protein as claimed in any one of claims 1 to 15, the process comprising coupling successive amino acid residues together and/or ligating oligo- and/or poly- peptides.

17. Synthetic or recombinant nucleic acid coding for a fusion protein as claimed in any one of claims 1 to 15.

18. Nucleic acid as claimed in claim 17, which is a vector.

19. A process for the preparation of nucleic acid as claimed in claim 17, the process comprising coupling successive nucleotides together and/or ligating oligo- and/or poly-nucleotides.

20. A cell or cell line transformed or transfected with a vector as claimed in claim 18.

21. A cell as claimed in claim 20, which is a yeast cell.

22. A yeast cell as claimed in claim 21 which is Pichia pastoris or Saccharomyces cerevisiae.

23. A cell as claimed in claim 20, which is a bacterial cell.

24. A bacterial cell as claimed in claim 23, which is Escherichia coli.

25. A pharmaceutical composition comprising one or more compounds as claimed in any one of claims 1 to 15 and a pharmaceutically or veterinarily acceptable carrier.

26. A method for the treatment or prophylaxis of thrombotic disease, the method comprising the administration of an effective, non-toxic amount of a fusion protein as claimed in any one of claims 1 to 15.

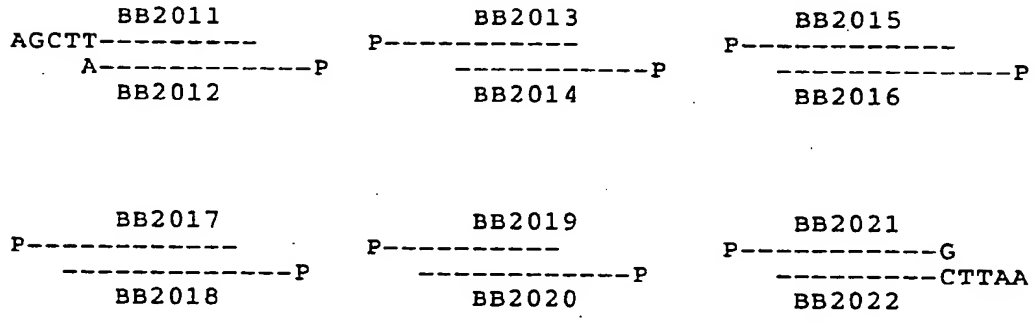
27. A proteinaceous compound as claimed in any one of claims 1 to 15 for use in human or veterinary medicine.

28. The use of a fusion protein as claimed in any one of claims 1 to 15 in the preparation of a thombolytic and/or antithrombotic agent.

FIG. 1.

SUMMARY OF ASSEMBLY PROCEDURE

The kinased oligomers were annealed in pairs. The oligomers BB2011 and BB2020 were not kinased to prevent multimerization.



p=5'phosphate

SUBSTITUTE SHEET

FIG. 2.

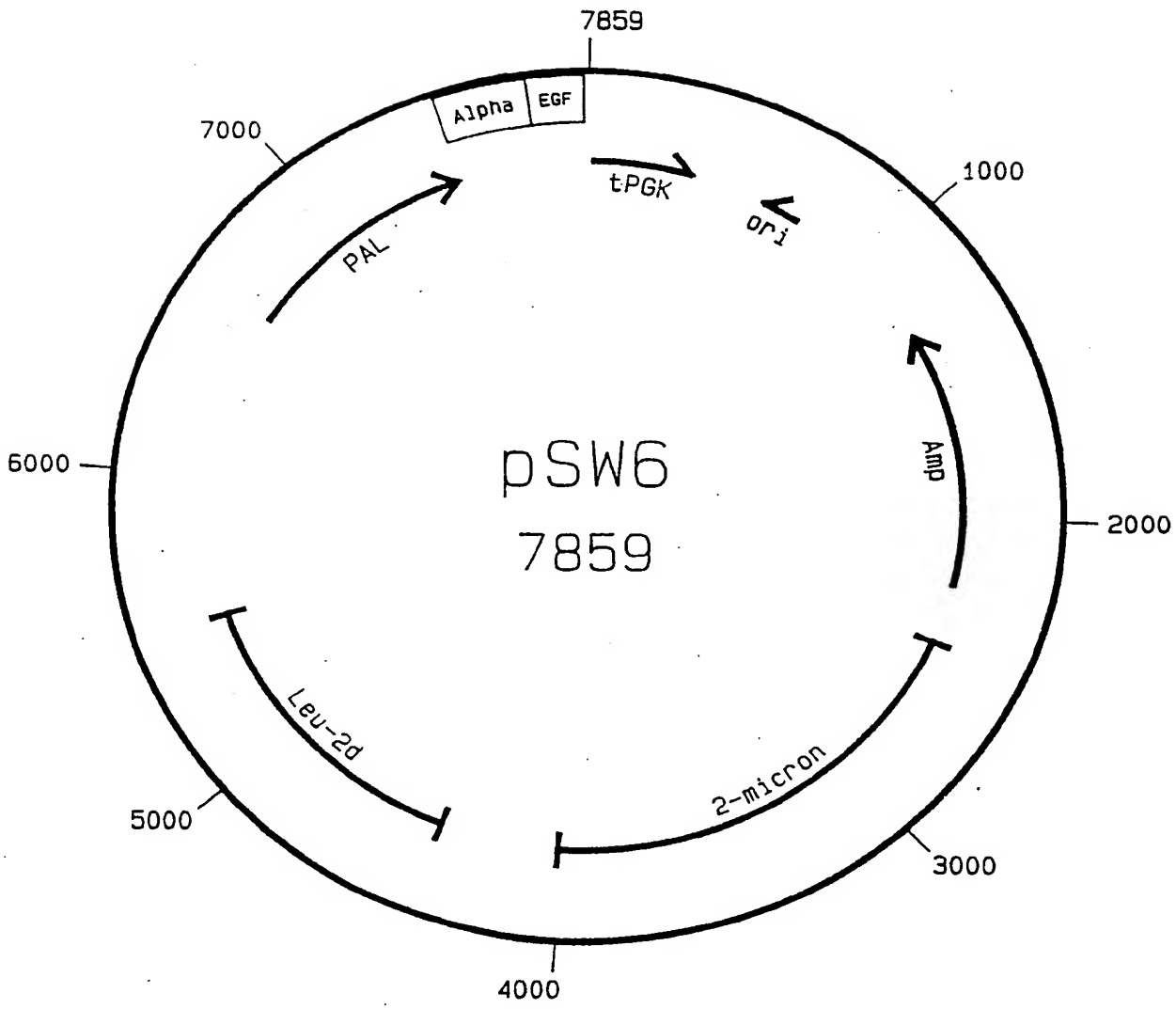


FIG. 3.

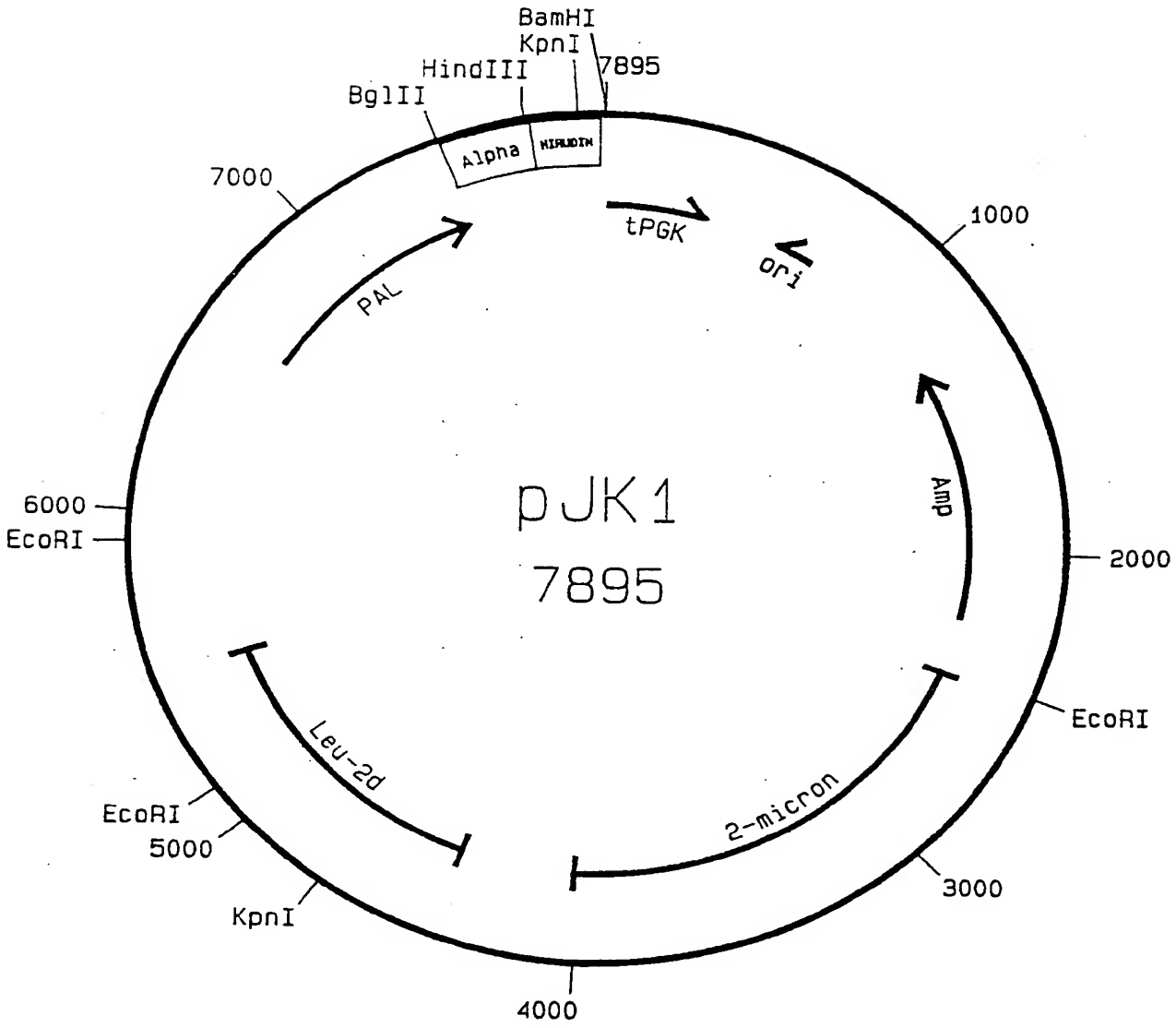
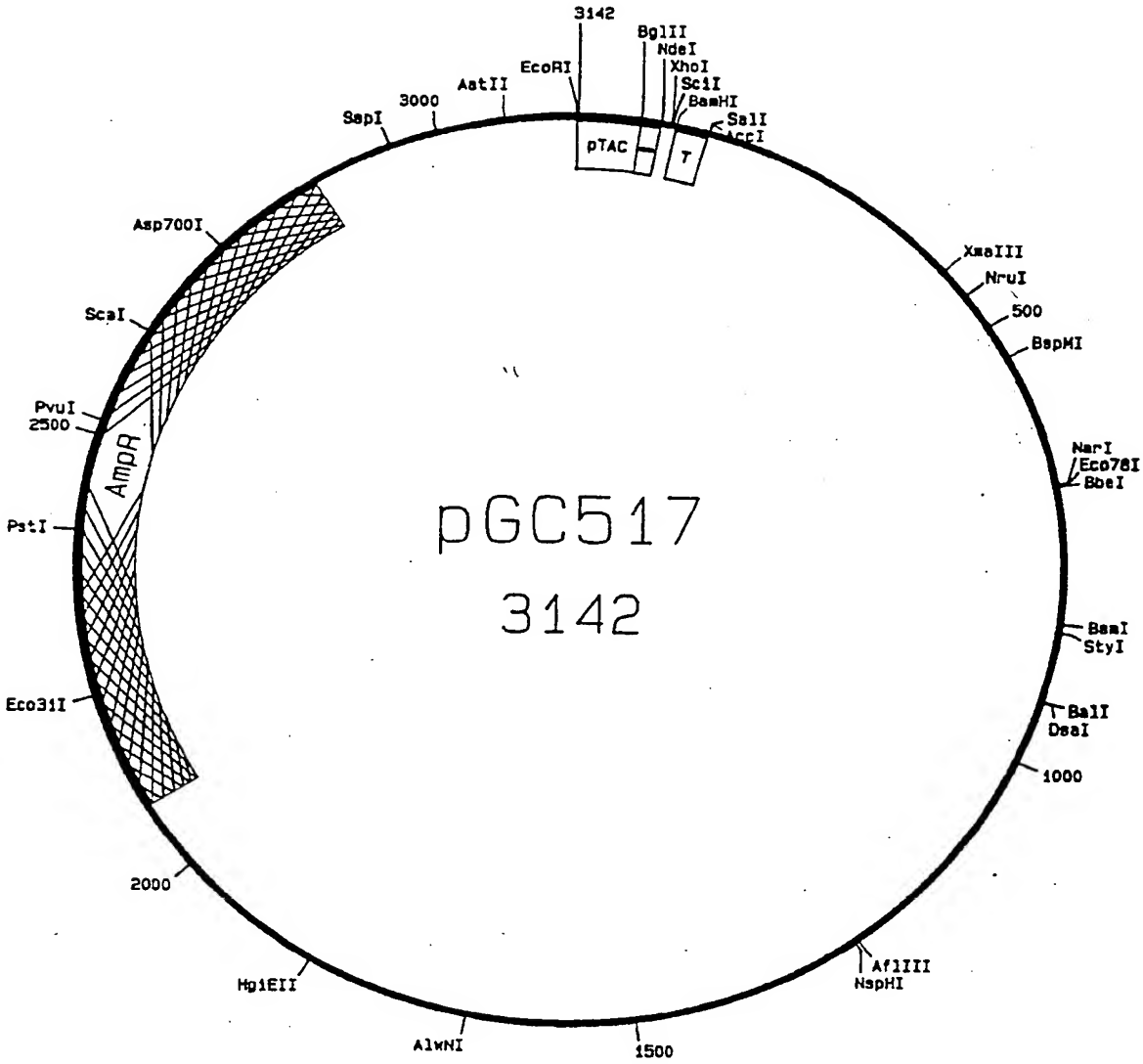
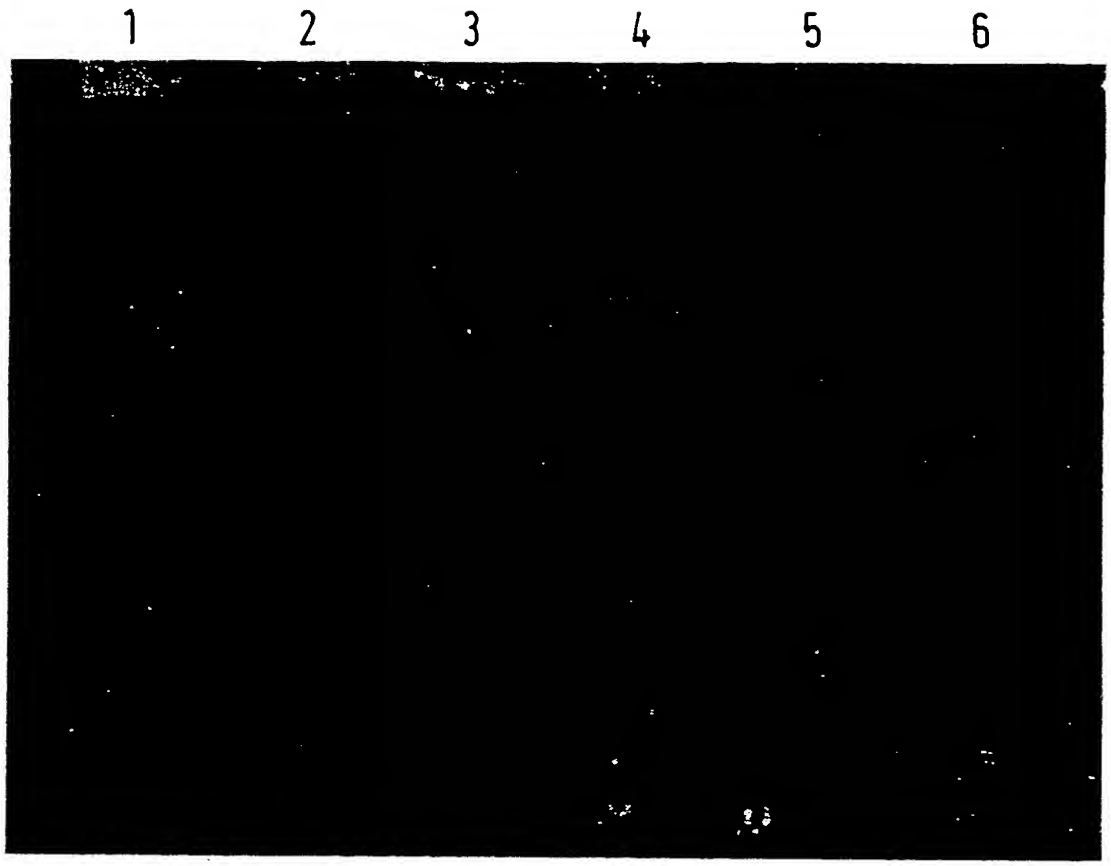


FIG. 4.



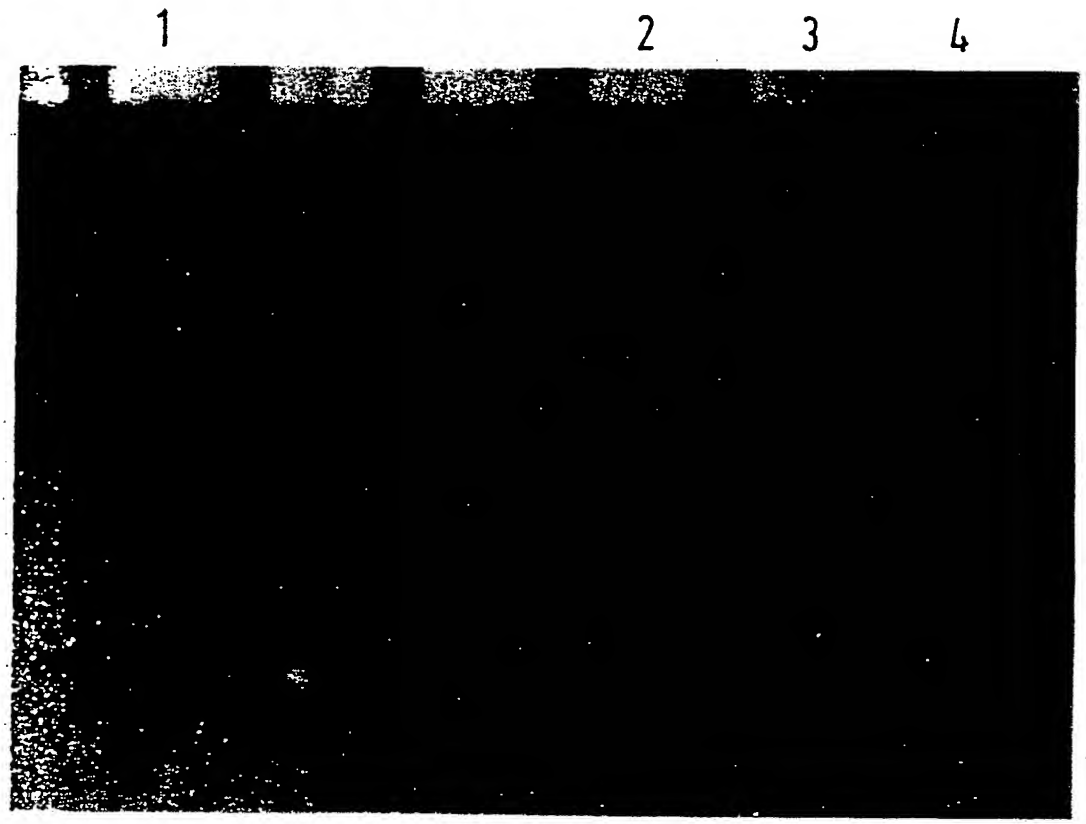
SUBSTITUTE SHEET

FIG. 5.



Zymograph of E. coli strains expressing streptokinase activity.
Lane 1. E. coli JM103 pKJ2 uninduced. Lane 2. E. coli JM103 pKJ2 IPTG induced. Lane 3. E. coli HW1110 pLGC1 uninduced. Lane 4. E. coli HW1110 pLGC1 IPTG induced. Lane 5. E. coli HW1110 pLGC2 uninduced. Lane 6. E. coli HW1110 pLGC2 IPTG induced.

FIG. 6.



Zymograph of in vitro cleavage of the thrombin cleavable Streptokinase-streptokinase molecule by thrombin.
Lane 1. Streptokinase. Lane 2. 15-40% cut containing high molecular weight streptokinase activity, no thrombin. Lane 3, as 2 but 0.5 U/ml thrombin. Lane 4, as 2 but 5 U/ml thrombin.

INTERNATIONAL SEARCH REPORT

International Application No PCT/GB 90/01911

I. CLASSIFICATI N OF SUBJECT MATTER (if several classification symbols apply, indicate all) [6]

According to International Patent Classification (IPC) or to both National Classification and IPC

IPC[5]: C 12 N 15/58, 15/62, 9/70, C 07 K 7/10, C 12 N 15/15, 1/19, 1/21, A 61 K 37/64, 37/54

II. FIELDS SEARCHED

Minimum Documentation Searched [7]

Classification System	Classification Symbols
IPC[5]	C 12 N, C 07 K, A 61 K

Documentation Searched other than Minimum Documentation to the Extent that such Documents are Included in the Fields Searched [8]

III. DOCUMENTS CONSIDERED TO BE RELEVANT [9]

Category *	Citation of Document, [11] with indication, where appropriate, of the relevant passages [12]	Relevant to Claim No. [13]
X	EP, A, 0296413 (HOECHST) 28 December 1988 see example 6	1,7,8,12, 16-20,25, 27,28
Y		15
Y	EP, A, 0323149 (ELI LILLY) 5 July 1989 see page 13, lines 19-30; page 16, lines 5-13	15
X	EP, A, 0292009 (ZYMOGENETICS) 23 November 1988 see page 3, lines 1-31; page 3, lines 57-58; page 6, line 48 - page 7, line 8; page 8, line 45 - page 9, line 15; page 22, section D; page 7, line 54 - page 8, line 2, examples 8,10,11; page 26, lines 38-45	1,7,8,12, 16-25,27,28
Y		2

./.

* Special categories of cited documents: [10]

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

IV. CERTIFICATION

Date of the Actual Completion of the International Search	Date of Mailing of this International Search Report
10th March 1991	08. 05. 91
International Searching Authority	**Signature of Authorized Officer**
EUROPEAN PATENT OFFICE	Harald Wrsuberg

Form PCT/ISA/210 (second sheet) (January 1985)

III. DOCUMENTS CONSIDERED TO BE RELEVANT (CONTINUED FROM THE SECOND SHEET)

Category *	Citation of Document, ** with indication, where appropriate, of the relevant passages	Relevant to Claim No.
A	EP, A, 0227938 (HOECHST) 8 July 1987 see page 3, lines 25-32; examples 1,2,4	3,9-12,25, 27,28
A	Mol Gen Genet, volume 212, 1988. MGG Springer-Verlag, C. Klessen et al.: "Tripartite streptokinase gene fusion vectors for gram-positive and gram-negative procaryotes", pages 295-300 see the whole document	4
A	EP, A, 0330700 (SAGAMI) 6 September 1989 see page 3, line 52 - page 4, line 45; page 5, line 20 - page 6, line 57	1,12

FURTHER INFORMATION CONTINUED FR M THE SECOND SHEET

V. [X] OBSERVATIONS WHERE CERTAIN CLAIMS WERE FOUND UNSEARCHABLE [1]

This international search report has not been established in respect of certain claims under Article 17(2) (a) for the following reasons:

1. [X] Claim numbers 26..... because they relate to subject matter not required to be searched by this Authority, namely:

see PCT Rule 39.1(iv)

2. [ ] Claim numbers..... because they relate to parts of the international application that do not comply with the prescribed requirements to such an extent that no meaningful international search can be carried out, specifically:

3. [ ] Claim numbers....., because they are dependent claims and are not drafted in accordance with the second and third sentences of PCT Rule 6.4(a).

VI. [ ] OBSERVATIONS WHERE UNITY OF INVENTION IS LACKING [2]

This International Searching Authority found multiple inventions in this international application as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this international search report covers all searchable claims of the international application.

2. [ ] As only some of the required additional search fees were timely paid by the applicant, this international search report covers only those claims of the international application for which fees were paid, specifically claims:

3. [ ] No required additional search fees were timely paid by the applicant. Consequently, this international search report is restricted to the invention first mentioned in the claims; it is covered by claim numbers:

4. [ ] As all searchable claims could be searched without effort justifying an additional fee, the International Searching Authority did not invite payment of any additional fee.

Remark on Protest

[ ] The additional search fees were accompanied by applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

ANNEX TO THE INTERNATIONAL SEARCH REPORT ON INTERNATIONAL PATENT APPLICATION NO. GB 9001911 SA 42783

This annex lists the patent family members relating to the patent documents cited in the above-mentioned international search report. The members are as contained in the European Patent Office EDP file on 23/04/91
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A- 0296413	28-12-88	JP-A- 1085096	30-03-89
EP-A- 0323149	05-07-89	AU-A- 2732988	29-06-89
		JP-A- 2002376	08-01-90
EP-A- 0292009	23-11-88	AU-A- 1652888	24-11-88
		JP-A- 1063379	09-03-89
EP-A- 0297882	04-01-89	AU-A- 1853388	19-01-89
		JP-A- 1080287	27-03-89
EP-A- 0207402	07-01-87	DE-A- 3523701	08-01-87
		JP-A- 62014795	23-01-87
EP-A- 0211299	25-02-87	DE-A- 3526995	05-02-87
		AU-B- 595221	29-03-90
		AU-A- 6056586	29-01-87
		JP-A- 62029600	07-02-87
EP-A- 0157235	09-10-85	DE-A- 3410437	26-09-85
		CA-A- 1230840	29-12-87
		JP-A- 60214897	28-10-85
WO-A- 9010081	07-09-90	AU-A- 5191790	26-09-90
WO-A- 8906239	13-07-89	EP-A- 0346500	20-12-89
EP-A- 0304013	22-02-89	JP-A- 2138995	28-05-90
EP-A- 0227938	08-07-87	DE-A- 3541856	04-06-87
		AU-B- 595262	29-03-90
		AU-A- 6569386	04-06-87
		JP-A- 62143696	26-06-87
EP-A- 0330700	06-09-89	WO-A- 8901513	23-02-89

EPO FORM P0479

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82